THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Analysis of partner proteins of MeCP2 and their relevance to Rett syndrome

Robert Ekiert

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

2012

Declaration

I declare that this thesis was composed by myself and the research presented is my own unless otherwise stated. This work has not been submitted for any other degree or personal qualification.

Robert Ekiert
October 2012

Table of contents

List of Figures

List of Tables

Acknowledgements

I first considered working for Adrian in the summer of 2005, while trying to find a better job than in a Fish & Chips in the suburbs of Edinburgh. I did not have courage to approach him then, so I learned all about preparing the full Scottish breakfast instead. Luckily, two years later the Wellcome Trust gave me a chance to do a rotation project in Adrian's lab. Although it wasn't very successful, I had lots of fun and I realised it was a great place to do science. The next four years proved I wasn't wrong and through the hurdles of my PhD project I learned a lot and enjoyed the whole experience. I am very grateful for having been able to work in Adrian's lab.

This thesis would not be possible without a bunch of other people that I have a great pleasure to know. Big thanks go to:

Matt, for a very fruitful collaboration, sharing the scientific ideas and resources and for proof-reading large parts of my thesis. I really enjoyed the endless discussions about paper reviewing process, good and bad science, but also about football and refereeing by Howard Webb. I still owe you 1½ dinners!

Jim, for always having a good word in scientific, career and life advice. Apart from your great help with mouse brains, I am very grateful for a chance to feel a bit Scottish in your kilt. You have to come to sunny Krakow to help me with pineapple and aubergine harvest!

Christine, for organising the best lab retreats in the world and for keeping Adrian on track. Without you the lab would be a disaster. Thanks for sorting all the issues with my money, conference travel and approaching deadlines and for not binning most of my lunches!

Dina, for reminding me every day that I should be writing instead of doing more experiments. How many times I wished I had listened to you!

Special thanks to the best immunohistochemistry teacher, Hélène, for constantly standing over my shoulder. Good luck in your new life role.

Sabine, for great teamwork in the mouse house and fun in sorting neurons. I'm really glad you were not deported for your issues with banks!

Sarah Young showed me that supervising a student, although demanding, can be fun and rewarding. Thanks for some good work which I used in this thesis.

All the former and present Bird lab members for being around at the various stages of my PhD, for keeping the lab a fun place to work and for sharing good times outside the campus too. Everyone helped in the final shape of this thesis as much as the constant supply of cakes and doughnuts did not help in my own shape. I can't name everyone here but you all know you will be missed!

Wellcome Trust for a generous scholarship and Adrian for allowing me to continue my research in his lab.

My parents and brother with his girls – for constant support and encouragement in my hours of doubt. After being so far from you for so long, I'm really happy I'm coming back and I hope it will be better than you say!

Finally, the biggest thank you goes to my beloved wife Marta. For her endless support, patience and understanding of my unpredictable working hours, forgetfulness and being mentally absent during the difficult weeks of writing. It should all be back to normal once it's finished and I hope we'll continue to be as happy as during last 12 years together!

The work in Adrian Bird's laboratory is funded by the Wellcome Trust and the Rett Syndrome Research Trust.

Abstract

Methyl-CpG binding protein 2 (MeCP2) was discovered as a protein binding to methylated DNA more than 20 years ago. It is very abundant in the brain and was shown to be able to repress transcription. The mutations in MeCP2 cause Rett syndrome, an autism-spectrum neurological disorder affecting girls. Yet, the exact role of MeCP2 in Rett disease, its function and mechanism of action are not fully elucidated. In order to shed some light on its role in the disease the aim of this project was to identify proteins interacting with MeCP2. Affinity purification of MeCP2 from mouse brains and mass spectrometry analysis revealed new interactions between MeCP2 and protein complexes. Detailed analysis confirmed the findings and narrowed down the top interactions to distinct regions of MeCP2. One of the domains interacts with identified NCoR/SMRT co-repressor complex and is mutated in many patients with Rett syndrome. *In vitro* assays proved that these mutations abolish the putative transcriptional repressor function of MeCP2. We propose a model in which Rett syndrome is caused by two types of mutations: either disrupting the interaction with DNA or affecting the interaction with the identified complex, which has an effect on the global state of chromatin. The presented findings can help to develop new therapies for Rett syndrome in the future.

Abbreviations

aa	amino acids
AID	activation-induced deaminase
APOBEC	apolipoprotein B mRNA-editing enzyme complex
Aprt	adenine phosphoribosyltransferase
ATRX	α thalassemia/mental retardation syndrome X-linked protein
BDNF	brain-derived neurotrophic factor
CAS	cellular apoptosis susceptibility protein
CK2α	casein kinase 2 α
CaMKIIα	calcium/calmodulin-dependent protein kinase II α
CDKL5	cyclin-dependent kinase-like 5
CGI	CpG island
CHD4	chromodomain helicase DNA binding protein 4
ChIP	chromatin immunoprecipitation
ChIP-seq	ChIP followed by high-throughput sequencing
CMV	cytomegalovirus
CREB1	cyclic-AMP responsive element binding protein 1
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
ES	embryonic stem (cells)
EST	expressed sequence tags
FBP11	formin-binding protein 11
FBS	foetal bovine serum
FRAP	fluorescence recovery after photobleaching
GFAP	glial fibrillary acidic protein
GFP	green fluorescent protein
GPS2	G protein pathway suppressor 2
HAT	histone acetyltransferase
HIPK2	homeodomain-interacting protein kinase 2
HDAC	histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HP1	heterochromatin protein 1
HRP	horseradish peroxidase
HYPC	huntingtin-interacting protein C
IP	immunoprecipitation
iPSC	induced pluripotent stem cells
kDa	kilodaltons
KO	knock-out
KPNA3	karyopherin α3 / importin subunit α3
KPNA4	karyopherin α4 / importin subunit α4

LANA	latency-associated nuclear antigen
LIF	leukaemia inhibitory factor
LINE	long interspersed nuclear element
MAP2	microtubule-associated protein 2
MeCP1, 2	methyl-CpG binding protein 1, 2
MBD	methyl-CpG binding domain
MBD1, 2, *etc*.	methyl-CpG binding domain protein 1, 2, *etc.*
MS	mass spectrometry
NCoR	nuclear receptor co-repressor
NLS	nuclear localisation signal
NMR	nuclear magnetic resonance
NP95	nuclear protein 95
NuRD	nucleosome remodelling and histone deacetylase complex
NURF	nucleosome remodelling factor complex
PAGE	polyacrylamide gel electrophoresis
PCH	pericentromeric heterochromatin
PCR	polymerase chain reaction
PBS	phosphate buffered saline
PTM	post-translational modifications
PWWP	proline-tryptophan- tryptophan-proline core domain
RA	retinoic acid
RBBP4	retinoblastoma binding protein 4
rpm	revolutions per minute
RT-qPCR	reverse transcription quantitative PCR
RTT	Rett syndrome
SAXS	small angle X-ray scattering
SMARC	SWI/SNF related, matrix associated, actin dependent regulator of chromatin
SDS	sodium dodecyl sulphate
SELEX	systematic evolution of ligands by exponential enrichment
SILAC	stable isotope labelling by amino acids in cell culture
SMRT	silencing mediator for retinoid or thyroid-hormone receptors
SUMO	small ubiquitin-like modifier
TAE	Tris-acetate/EDTA
TBL1	transducin β-like 1
TBLR1	transducin β-like 1 related protein
TBS	Tris buffered saline
TET	ten-eleven translocation
TFA	trifluoroacetic acid
TFIIB	transcription factor IIB
TK	thymidine kinase
TRD	transcriptional repression domain
TSA	trichostatin A
UAS	upstream activating sequence
UHRF1	ubiquitin-like, containing PHD and RING finger domains 1
UTR	untranslated region
YB-1	Y-box binding protein 1
YY1	yin yang 1

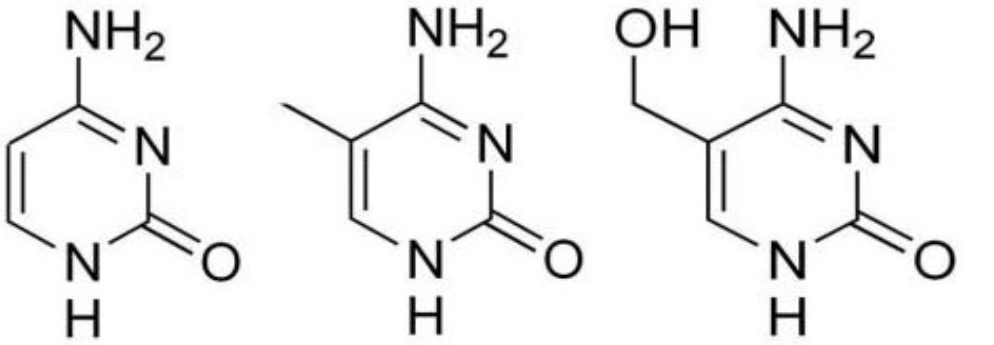
Chapter 1: Introduction

1.1 Modifications of DNA

The genetic information in all cellular organisms is encoded in deoxyribonucleic acid (DNA). It is build of only four nucleotides: adenine, guanine, thymine and cytosine, with backbones made of sugars and phosphate groups. In the DNA double helix adenine complementarily pairs with thymine and guanine with cytosine, allowing faithful replication of the genome and its inheritance through cell divisions (Watson and Crick, 1953). The four building blocks of DNA are sufficient to encode the information needed to build any protein from twenty standard amino acids by usage of codons consisting of nucleotide triplets. Despite the redundancy of such a genetic code, the DNA bases are additionally found to be modified by chemical addition of covalently bound moieties, expanding the regulatory capacity of the genome. For example, in prokaryotes adenine is frequently methylated on nitrogen-6 and cytosine exists in methylated states on carbon-5 and nitrogen-4 (Ehrlich et al., 1987). Bacterial DNA methylation serves three purposes: (1) it allows distinguishing between self and intruder DNA, permitting the defence mechanism by methyl-sensitive restriction digestion degrading only the non-host genome (reviewed in: Wilson and Murray, 1991); (2) it ensures fidelity of the genome replication by DNA mismatch repair machinery operating on non-methylated newly synthesised strand (Laengle-Rouault et al., 1986); and (3) it regulates transcription of genes (reviewed in: Palmer and Marinus, 1994).

1.1.1 DNA methylation and hydroxymethylation in eukaryotes

In eukaryotes the DNA modifications are found exclusively on cytosine in the form of 5-methylcytosine and its derivative 5-hydroxymethylcytosine (Fig. 1.1). Two additional forms of cytosine detected in mouse cells are 5-formylcytosine and 5-carboxylcytosine (Ito et al., 2011), but as yet they were studied to a much lower extent than the methylated and hydroxymethyalted versions.



cytosine 5-methylcytosine 5-hydroxymethylcytosine

Figure 1.1. Chemical structure of cytosine nucleotide and its modified forms.

The occurrence, distribution and function of methylcytosine and its derivatives vary between the organisms, and are not always correlated with the course of evolution.

The lower eukaryotes such as budding and fission yeast *Sacharomyces cerevisae* and *Schizosaccharomyces pombe* contain no detectable DNA modifications (Antequera et al., 1984; Proffitt et al., 1984). In other fungi, *Sporotrichum dimorphosporum*, *Phycomyces blakesleeanus* or *Neurospora crassa* 5-methylcytosine was found at levels ranging from 0.2% up to 3% of all cytosines (Antequera et al., 1984; Selker et al., 2003).

Invertebrates also show variable amounts of DNA methylation: from undetectable in nematode *Caenorhabditis elegans*, through trace levels in fruit fly *Drosophila melanogaster* to long stretches of methylated cytosines in sea urchin *Echinus esculentus* and sea squirt *Ciona intestinalis* (Bird et al., 1979; Antequera et al., 1984; Simmen et al., 1999; Lyko et al., 2000).

With no exceptions found to date, the genomes of all vertebrates are globally methylated (Bird and Taggart, 1980; Tweedie and Charlton, 1997). The 5-methylcytosine mark occurs almost exclusively symmetrically in the context of a palindromic dinucleotide, CpG. CpGs are broadly distributed in the genome and are methylated in 60-90% (Tweedie and Charlton, 1997). Low levels of non-CpG methylation (CpA, CpT and CpC) were found in embryonic stem (ES) cells and induced pluripotent stem cells (iPSC) but are nearly absent in somatic cells (Ramsahoye et al., 2000; Imamura et al., 2005; Lister et al., 2009; Ziller et al., 2011).

Interestingly, in the vertebrate bulk genome the CpG sequences are present at only one-fifth of the expected frequency (Bird, 1980). There are, however, regions of the genome rich in guanine and cytosine where the CpG dinucleotides are much more frequent. These ~1 kb long DNA stretches constitute less than 2% of the mouse genome, are predominantly unmethylated and were named CpG islands (CGIs; Bird et al., 1985). Based on genome-wide sequencing the numbers of CGIs in mouse and human genomes were estimated to be approximately 23 000 and 25 500, respectively (Illingworth et al., 2008, 2010). About 60% of CGIs are associated with promoters of known genes. Moreover, the remaining CGIs have also been linked to active transcription (Illingworth et al., 2010; Maunakea et al., 2010). These facts suggest possible mechanism of the origin of CpG islands. 5-methylcytosine is a potential mutagen as it can spontaneously deaminate to thymine. A resulting T:G mismatch with the other DNA strand can not be efficiently repaired (reviewed in: Bird, 1986). That led to under-representation of CpGs in the bulk genome. However, when the active fragments of genome were devoid of cytosine methylation, the CpG sequences were preserved during evolution at expected frequency. It also suggests that the DNA methylation in vertebrates may be associated with transcriptional repression.

In plants the genome contains relatively high levels of methylation, between 5-25% of all cytosines depending on species (reviewed in: Rangwala and Richards, 2004). The methyl mark occurs in the CpG context, as well as CpNpG and asymmetric CpHpH, where N is any nucleotide and H is any nucleotide but guanine. Genome-wide studies in *Arabidopsis thaliana* revealed that methylated cytosine is excluded from gene promoters, covers transposons and silent heterochromatin, but unlike in mammals, only 20% of the gene bodies are methylated (Zilberman et al., 2007).

Another DNA modification, 5-hydroxymethylcytosine (Fig. 1.1) has gained a vast scientific attention in the last years. Although it was known for more than a half a century in bacteriophages and viruses (Wyatt and Cohen, 1952, 1953) only recently it has been discovered to be naturally present in Purkinje and embryonic stem cells in mice, at the level up to 0.6% of all nucleotides (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). It arises from 5-methylcytosine by conversion catalysed by

ten-eleven translocation (TET) proteins. It has been proposed to be an intermediate product of DNA de-methylation (both in passive or the putative active process), but may as well play a role on its own (Tahiliani et al., 2009). Using various techniques for genome-wide mapping of the 5-hydroxymethylcytosine mark it has been found to be associated with gene enhancers, transcriptional start sites, as well as with gene bodies (Pastor et al., 2011; Stroud et al., 2011; Williams et al., 2011). Its proposed roles in gene expression regulation or DNA demethylation are still awaiting confirmation and detailed analysis.

Because of the high abundance of 5-methylcytosine in mammals and the various functions associated with it (described below), it is sometimes called 'the fifth base'. Whether 5-hydroxymethylcytosine should be referred to as 'the sixth base' is still uncertain. The recent and coming years of increased research on this form of cytosine give hope for obtaining an answer to this question in the near future.

1.1.2 Processing of DNA modifications

The mechanism of deposition of the methylation mark on cytosines is known and has been well studied. In mammals *de novo* DNA methylation is catalysed by DNA methyltransferase enzymes, DNMT3a and DNMT3b (Okano et al., 1999). In somatic cells methylation patterns are relatively stable and are heritable through cell divisions. During DNA replication in the S-phase of cell cycle, the 5-methylcytosine mark is copied onto the newly synthesised strand by a so-called maintenance DNA methyltransferase, DNMT1 (for a comprehensive review see: Goll and Bestor, 2005). Conversely, how the genomic methylation is removed is still not known. Demethylation occurs globally twice during development – in primordial germ cells and in the preimplantation embryos (reviewed in: Reik et al., 2001). According to one theory it may be a passive mechanism, when during replication the mark is not copied then in subsequent cell divisions the methylated cytosines are simply diluted away. Another possibility is that 5-methylcytosine is converted back to cytosine in an active process. The rapid demethylation of paternal genome in mouse zygote has been observed, however the mechanism and the enzymes involved in this pathway have not been found (Oswald et al., 2000; Santos et al., 2002). There is some indication that TET proteins may perform that role by converting 5-methylcytosine

to its hydroxylated form. 5-hydroxymethylcytosine can be subsequently processed to 5-hydroxyuracil by AID/APOBEC (activation-induced deaminase / apolipoprotein B mRNA-editing enzyme complex) and removed by base excision repair mechanism (Guo et al., 2011). However, the ability of AID/APOBEC to perform this convertion has been recently questioned (Nabel et al., 2012). Another hypothetical model of active de-methylation came with discovery that 5-hydroxymethylcytosine can be processed by TET enzymes to 5-formylcytosine and 5-carboxylcytosine (Ito et al., 2011). These forms of cytosine can be removed by thymine DNA glycosylase, allowing the base excision repair mechanism to fill in the missing cytosine (Maiti and Drohat, 2011). However, the genome-wide generation of nucleotide excisions does not seem like a feasible mechanism for de-methylation as it is burdened with risk of introducing mutations. Moreover, the immunofluorescent staining of 5-hydroxymethylcytosine in early stage zygote revealed that it is gradually reduced during cell divisions supporting the dilution hypothesis (Inoue and Zhang, 2011). The search for the exact de-methylation mechanism continues.

1.1.3 The roles of DNA methylation

The loss-of-function mutation of *Dnmt1* in mice, as well as double deletion of *Dnmt3a* and *Dnmt3b,* were embryonic lethal, suggesting that the establishment and maintenance of DNA methylation are crucial for the mouse development (Li et al., 1992; Okano et al., 1999).

The most studied function of DNA methylation is repression of transcription. It has been observed that when cells were treated with 5-azacytidine, a nucleoside analogue causing hypomethylation of newly synthesised DNA (Jones and Taylor, 1980), the expression of many genes was activated (Mohandas et al., 1981; Venolia et al., 1982). Other evidence for methylation-mediated transcriptional silencing was provided by introducing the adenine phosphoribosyltransferase (aprt) gene into *aprt*$^-$ mouse cells. When the transgene was *in vitro* methylated, it was not expressed and could not rescue the cells phenotype in contrast to its unmethylated version (Stein et al., 1982). There were many further reports based on methylated reporter genes and supporting the idea of CpG methylation leading to gene repression (Buschhausen

et al., 1985; Boyes and Bird, 1991; Siegfried et al., 1999). More recent studies employing genome-wide sequencing techniques also support this notion (Weber et al., 2007; Illingworth et al., 2008).

Another well-studied process involving DNA methylation is X chromosome inactivation. In mammals female genome contains two X chromosomes, whereas male only one. To compensate the dosage of X-linked genes between genders, one copy of female X has to be silenced (Lyon, 1961). The chromosome to be inactivated (X_i) is chosen randomly in each cell during embryogenesis and this choice is propagated through cell divisions. Silencing is mediated by coating the X_i by non-coding RNA *Xist* (reviewed in: Brockdorff, 2011). Methylation of CpG islands of gene promoters on X_i stabilises the inactive state and allows its heritability (Heard and Disteche, 2006). The early experiments with 5-azacytidine reactivating genes on the inactive X chromosome also supported the idea of the role of DNA methylation in dosage compensation in mammals (Mohandas et al., 1981; Venolia et al., 1982).

DNA methylation also has a function in genomic imprinting, where genes are expressed from only one allele determined by parental origin. The genomic sites exhibiting parent-specific methylation are called differentially methylated regions (DMR) and control the silencing of the locus (reviewed in: Edwards and Ferguson-Smith, 2007). Loss of DNA methylation at DMRs leads to activation of normally silent, imprinted loci (Li et al., 1993).

DNA methylation additionally plays an important role in cancer development. Although cancer cells are generally hypomethylated, the CGI promoters of many tumour suppressor genes gain methylation (Ehrlich, 2009). In this way genes involved in various cellular processes, such as DNA repair (hMLH1, MGMT, BRCA1), cell cycle (p16INK4b, Rb), cell adherence (E-cadherin), signalling (RASSFIA, APC) or apoptosis (TMS1), are silenced and this contributes to tumour progression (reviewed in: Esteller, 2007).

The mammalian genome consists of more than 35% repetitive sequences, such as L1 and Alu retrotransposons, which are relics of intragenomic parasites (Yoder et al., 1997). As the bulk genome is largely methylated the modified CpGs have been found in these sequences too. It has been proposed that the main function of DNA methylation is to protect the genome from retrotransposition events by

repressing transcription of these transposable elements (Yoder et al., 1997). However, there is not much evidence to support these arguments. In particular, in germ cells and during embryogenesis the genome, including retrotransposons, is demethylated. Moreover, genome-wide studies did not find the transposable elements as particularly targeted for CpG methylation in *Ciona intestinalis* (Suzuki et al., 2007) and in mammals (Rabinowicz et al., 2003).

Another possible function of the CpG methylation comes from the complexity of the mammalian genome. The methylation of bulk genome could reduce the 'transcriptional noise' from irrelevant, cryptic promoters, which could otherwise interfere with expression of important genes (Bird, 1995).

How exactly DNA methylation mediates the ascribed functions is still a topic of a debate. Theoretically, it is possible that the methyl residue on cytosine causes exclusion of DNA binding proteins which have affinity for the unmethylated sequence. It was shown to be the case for many proteins, such as c-Myc (Prendergast et al., 1991), Cfp1 (Thomson et al., 2010), or CTCF (Hark et al., 2000). Another possibility is that there are specialised proteins recognising and binding to DNA containing methylated CpGs. This notion was supported by observed protection of methylated DNA from restriction digestion in mammalian nuclei (Antequera et al., 1989).

1.2 Methyl-CpG binding proteins

The first search for proteins binding to methylated DNA revealed an activity named methyl-CpG binding protein 1 (MeCP1; Meehan et al., 1989). The protein identities of this large complex remained unknown for several years. Its discovery was however followed by identification of another methyl-CpG binding activity, named MeCP2, which was proven to be different from MeCP1 by size, DNA binding preference, biochemical properties and abundance in cells (Lewis et al., 1992; Meehan et al., 1992). It has been sequenced as one protein and the *Mecp2* gene was cloned from rat brain (Lewis et al., 1992). As MeCP2 protein is the main focus of this thesis and it is described in detail in section 1.3.

1.2.1 MBD proteins

Deletion analysis of MeCP2 allowed identification of the minimal domain responsible for interaction with methylated cytosines, which was named methyl-CpG binding domain (MBD; Nan et al., 1993). The search for similar sequences in the expressed sequence tags database (dbEST) resulted in identification of a protein containing MBD domain 1 (PCM1; Cross et al., 1997), later re-named MBD1. Similar searches in mouse EST database found other genes: *Mbd1, Mbd2, Mbd3* and *Mbd4* (Hendrich and Bird, 1998), which were subsequently cloned and mapped (Hendrich et al., 1999a). All of the identified MBD proteins, with the exception of MBD3, were able to bind to methylated DNA *in vitro* and localised in the pericentromeric heterochromatin (PCH) foci *in vivo* (Hendrich and Bird, 1998). The PCH in mice consists of major satellite repeats, which are known to be rich in methylated sequences (Miller et al., 1974; Hörz et al., 1981).

Interestingly, even in cells with a mutated *Dnmt1* gene and in consequence >95% reduced levels of methylation, MBD1 was still localised in the PCH (Hendrich and Bird, 1998; Jørgensen et al., 2004). The non-methyl CpG binding was confirmed *in vitro* and was attributed to one of the three cystein-rich CxxC domains present in the MBD1 structure (Fig. 1.2A). MBD1 also contains a transcriptional repression domain (TRD) and was shown to repress transcription of both methylated and non-methylated reporter genes (Fujita et al., 1999; Ng et al., 2000; Jørgensen et al., 2004).

MBD2 binds to a single methylated CpG without other sequence specificity (Ng et al., 1999; Klose et al., 2005). It has been found to be always associated with the nucleosome remodelling and histone deacetylase (NuRD) co-repressor complex (Le Guezennec et al., 2006). This complex was identified as a previously described methyl-CpG binding activity, MeCP1. It is able to repress transcription in a histone-deacetylase dependent manner (Ng et al., 1999).

Although mammalian MBD3 is closely related to MBD2 with 70% of their sequences identical (Hendrich and Tweedie, 2003), it is not capable of binding to methylated DNA due to a missense mutation in the MBD domain (Hendrich and Bird, 1998; Fraga, 2003). Its function is unknown, but it has been reported to be a part of NuRD repressor complex distinct from MBD2/NuRD (Le Guezennec et al.,

2006). There is also a preliminary indication that its role depends on binding to 5-hydroxymethylcytosine (Yildirim et al., 2011).

MBD4 was demonstrated to bind preferentially to methyl-CpG : TpG mismatches and to remove thymine or uracil *in vitro* (Hendrich et al., 1999b). The glycosylase domain responsible for this activity has been mapped to the C-terminus of the protein and is unique to MBD4 in the MBD protein family (Fig. 1.2A). As the C→T transitions are the primary product of deamination of 5-methylcytosine, it was suggested that MBD4 plays a role in repairing such mutations. MBD4 was additionally shown to be able to repress transcription of reporter genes through interaction with Sin3A and histone deacetylase HDAC1 (Kondo et al., 2005).

By less stringent bioinformatic sequence similarity searches two additional proteins containing a homologue of methyl-CpG binding domain have been discovered and named MBD5 and MBD6 (Hendrich and Tweedie, 2003; Roloff et al., 2003; Laget et al., 2010). Their MBD domains, however, include a 9-amino acid deletion at the region conserved between other MBDs. MBD5 contains also a proline-tryptophan-tryptohan-proline core domain (PWWP) that may bind to histones (Fig. 1.2A; Laget et al., 2010). The properties and function of these new MBD family members have not been extensively studied, but it was reported that, although they localise in the heterochromatic foci in cells, both MBD5 and MBD6 are incapable of binding to methylated DNA *in vitro* (Laget et al., 2010).

1.2.2 Kaiso and other proteins binding to methylated DNA

There is another family of proteins that binds preferentially to methylated DNA, but lacks sequence similarity to MBD proteins (Fig. 1.2B). The founder member of the family, Kaiso, was shown to interact specifically with doubly methylated sequence CGCG, through the conserved zinc fingers present in its structure (Prokhortchouk et al., 2001; Daniel et al., 2002; Buck-Koehntop et al., 2012). It is also able to repress transcription by interacting with the nuclear receptor co-repressor (NCoR) complex (Yoon et al., 2003). Two other members of the family, Kaiso-like proteins ZBTB4 and ZBTB38 require only a single methylated CpG to bind to DNA and are capable of repressing transcription as well (Filion et al., 2006).

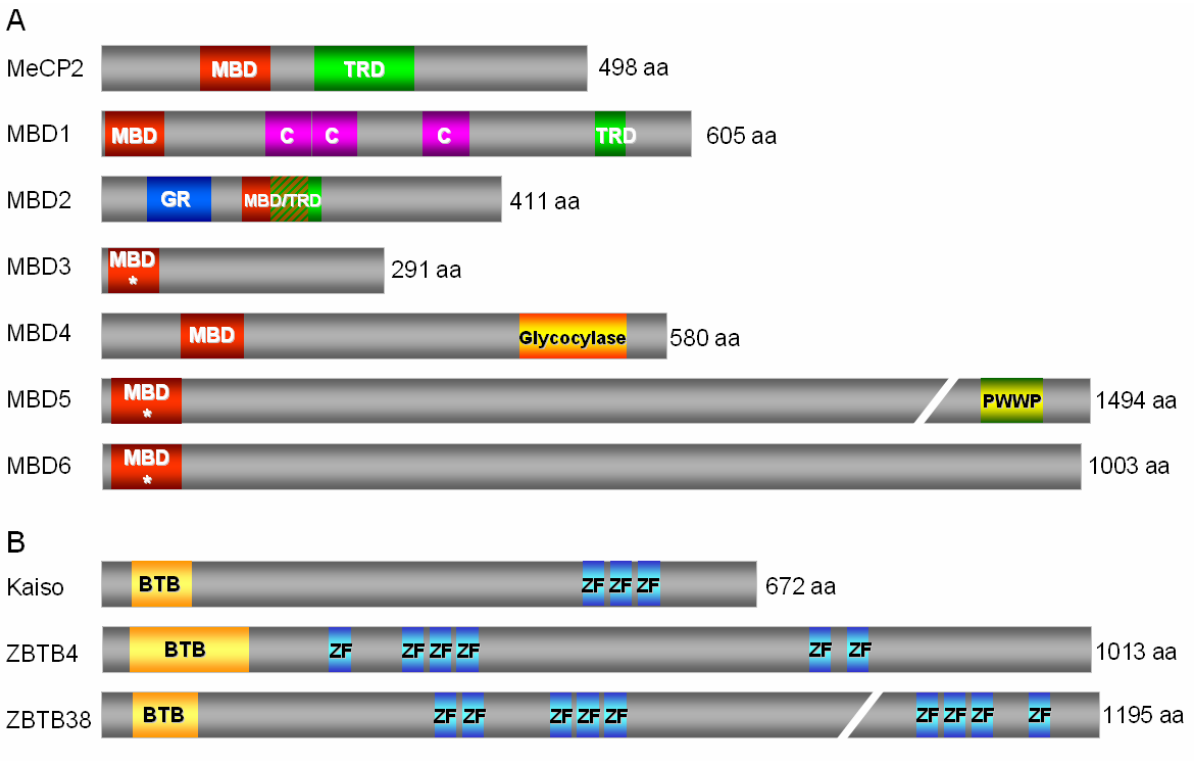


Figure 1.2. Methyl-CpG binding proteins. Schematic representation of proteins belonging to (A) MBD and (B) KAISO protein families. Where multiple isoforms of a protein exist, the longest version of human protein was depicted, with size in amino acid stated on the right. The domain coordinates were retrieved from PROSITE database (Sigrist et al., 2010) or taken from the literature. MBD – methyl-CpG binding domain, TRD – transcriptional repression domain, C – CxxC domain, GR – Glycine/Arginine repeats, PWWP – Proline-Tryptophan-Tryptophan-Proline core domain, BTB – BTB/POZ domain, ZF – zinc finger.

There are additional proteins postulated to recognize and bind to methylated DNA, but have not been studied in details. UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1; also known as nuclear protein 95, NP95) may play a role in maintenance of DNA methylation as it binds to hemimethylated DNA and is able to recruit DNMT1 (Bostick et al., 2007; Sharif et al., 2007). Interesting approach based on SILAC method (stable isotope labelling by amino acid in cell culture) and aimed at identification proteins binding to methylated CpGs identified most of the proteins described above and additionally detected RBP-J (Bartels et al., 2011). Its binding to methylated DNA was shown to be sequence-specific but its functional consequence has not been investigated (Bartels et al., 2011).

1.2.3 Mouse models depleted of methyl-CpG binding proteins

The best method of assessing the role and importance of a gene is to generate knock-out animals and evaluate their phenotype (Jackson, 2001). The importance of DNA methylation for mouse development was demonstrated by embryonic lethality of mice with mutated *Dnmt1* or *Dnmt3a* and *Dnmt3b* (Li et al., 1992; Okano et al., 1999). Interestingly, deletion of different methyl-CpG binding proteins gave mixed outcomes suggesting their, at least partial, redundancy or a role of DNA methylation independent of proteins binding to it. The most severe effects come from deletion of *Mecp2*, leading to a neurological disorder resembling closely Rett syndrome in humans (Chen et al., 2001; Guy et al., 2001; see section 1.4). A much weaker phenotype was exhibited by *Mbd1*-null animals, with mild deficits in neurogenesis, memory formation, long-term potentiation and symptoms of autism-like behaviour (Zhao et al., 2003; Allan et al., 2008).

Mbd2 knock-out mice are viable and fertile and have normal appearance with only mild phenotypes. The $Mbd2^{-/-}$ mothers fail to properly nurture their pups, who in consequence are significantly smaller (Hendrich et al., 2001). Deficiency of MBD2 also leads to subtle differences in gene expression patterns, as tested in T-helper cells (Hutchins et al., 2002) and in the colon (Berger et al., 2007).

The knock-out of *Mbd3* was embryonic lethal (Hendrich et al., 2001). The function of the protein is unknown, but it was shown to play a role in ES cell differentiation by down-regulation of pluripotency markers (Kaji et al., 2007).

$Mbd4^{-/-}$ mice showed 3.3-fold increase in frequency of CG → TA transitions at CpG sites (Millar et al., 2002). Moreover, when crossed with cancer-susceptible $Apc^{min/+}$ mice, accelerated tumour formation and reduced animal survival have been observed (Millar et al., 2002). The results implicate a role of MBD4 in suppressing the mutability of methylated CpG loci.

Deletion of *Kaiso* in *Xenopus* resulted in severe developmental problems (Ruzov et al., 2004), but surprisingly caused no overt phenotype in mice (Prokhortchouk et al., 2006). Interestingly, the effect of crossing *Kaiso*-null with $Apc^{min/+}$ mice was a delayed onset of tumour formation (Prokhortchouk et al., 2006).

The reason for existence of so many proteins able to bind to the same methylated sequences and repress transcription is unknown. The unique domains on each of the proteins suggest that they may have distinct interaction partners and therefore perform non-redundant functions. As different proteins exhibit variable expression levels across the organs it is possible that each of them is essential in different tissue. This notion, however, is not supported by the mutational analysis of MBD and Kaiso proteins in mice described above. Double and triple knock-outs of *Mbd2, MeCP2* and *Kaiso* suggested that these proteins play non-overlapping roles in mouse development, but may be redundant to some extent in mature animals and neurons (Martín Caballero et al., 2009). There is also an indication that the sequence specificity of some of the MBD proteins is non-redundant and MeCP2, MBD1 and Kaiso possess their distinctive binding sites in the genome (Daniel et al., 2002; Klose et al., 2005; Clouaire et al., 2010). This idea remains to be confirmed *in vivo*, however some preliminary chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq) has been performed for MeCP2. Its global binding profile in the brain suggests that it interacts with any methylated CpG it can encounter (Skene et al., 2010). Additionally, the relationship between MBD and Kaiso-like proteins and the 5-hydroxymethylcytosine mark is not well established yet. The coming years of greatly enhanced research in this field should provide us with more understanding of the complex system of DNA methylation, hydroxymethylation and their mediators.

1.3 MeCP2

1.3.1 Gene and protein

Methyl-CpG binding protein 2 was discovered and described as a protein with affinity to methylated DNA in 1992 (Lewis et al. 1992; Meehan et al. 1992). In humans it is encoded by a single gene on the X chromosome in locus Xq28 (D'Esposito et al., 1996). MeCP2 protein was found to be expressed ubiquitously in human and mouse tissues, but is particularly abundant in brain, especially in neurons

(LaSalle et al., 2001; Shahbazian et al., 2002b). The quantification of MeCP2 expression in neuronal nuclei sorted from adult mouse brains revealed that there are approximately 16×10^6 molecules, compared to 2×10^6 in glial cells and 0.5×10^6 in liver (Skene et al., 2010). Such a high abundance, almost reaching the number of nucleosomes (estimated to be 30×10^6) suggests the fundamental role of MeCP2 in neurons.

MeCP2 occurs in two splicing variants that differ at the N-terminal end of the protein (Kriaucionis and Bird, 2004; Mnatzakanian et al., 2004). The mRNA of human MeCP2 variant e2 (also known as isoform 1, MeCP2A and MeCP2β) consists of 4 exons, with the translational start site located in exon 2, giving a protein of 486 amino acids (aa) and a predicted molecular weight of 52.4 kDa (Fig. 1.3). The second splicing variant, named MeCP2_e1 (also known as isoform 2, MeCP2B and MeCP2α) starts in exon 1 and omits exon 2, consists of 498 aa (53.5 kDa) and it was reported to be the most abundant isoform in mouse and human brains (Kriaucionis and Bird, 2004; Mnatzakanian et al., 2004). Both isoforms were expressed in all tested tissues and seem to function in the same way. They localise in the same regions of chromatin (Kriaucionis and Bird, 2004) and exhibit identical dynamics and mobility (Kumar et al., 2008). One isoform can successfully substitute the other in the *Xenopus laevis* embryo without any noticeable effects on its development (Stancheva et al., 2003). Also a very recent study showed that the phenotype of *Mecp2*-null mice is rescued by transgenic expression of MeCP2_e1 as well as MeCP2_e2 (Kerr et al., 2012).

Both isoforms contain two well-studied functional domains (Fig. 1.3). The methyl-CpG binding domain (MBD) is responsible for the interactions with methylated DNA (Nan et al., 1993) and transcriptional repression domain (TRD) is able to repress transcription (Nan et al., 1997). Alignment of MeCP2 sequence from *Danio rerio, Xenopus laevis, Rattus norvegicus, Mus musculus, Macaca fascicularis* and *Homo sapiens* revealed that MeCP2 is highly conserved from fish to humans, especially within the MBD and TRD regions (Thambirajah et al., 2009). Interestingly, multiple polyadenylation sites give rise to transcripts of various lengths, including an unusually long (10.2 kb) 3' untranslated region (3' UTR; Singh et al., 2008). Several

regions of this UTR are conserved between human and mouse and were suggested to be involved in the regulation of expression levels of MeCP2 (Coy et al., 1999). The exact mechanism of this regulation is unknown, however two microRNAs, miR-132 and miR-212 were reported to bind 3' UTR of MeCP2 and post-transcriptionally diminish its translation (Klein et al., 2007; Wada et al., 2010).

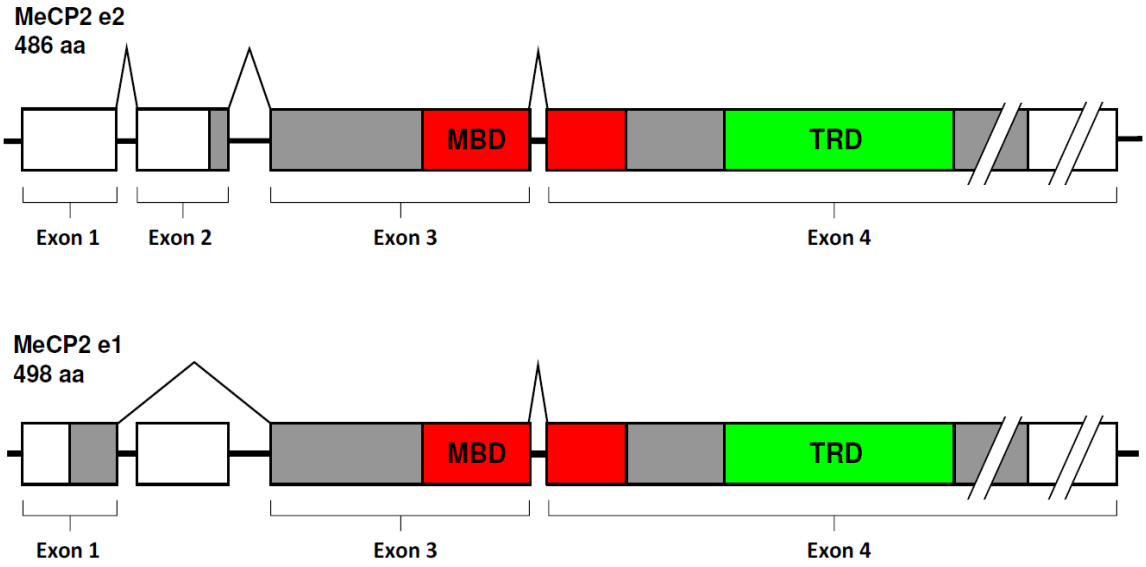


Figure 1.3. MeCP2 splicing variants. White rectangles depict the non-coding regions of *MECP2* exons. Translated protein is marked in grey. Conserved domains are highlighted in red and green and their function is described in subsequent sections. MBD – methyl-CpG binding domain, TRD – transcriptional repression domain. The sizes of proteins in amino acids (aa) correspond to human MeCP2. Introns are not represented to scale.

Importantly, mutations in the *MECP2* gene in humans cause Rett syndrome (RTT), a neurological disease, which is described in detail in section 1.4 of this chapter. *Mecp2* gene was knocked-out in mice giving phenotype similar to RTT, which is described later.

Despite its calculated mass MeCP2 migrates on an electrophoretic gel with an apparent size of about 80 kDa. In gel filtration studies it eluted with an apparent mass of 400-500 kDa (Klose and Bird, 2004). Analytical centrifugation revealed that these discrepancies were not due to self-association of MeCP2, nor because of interactions

with nucleic acids or other proteins (Klose and Bird, 2004). The explanation suggested by authors was that MeCP2 monomers exhibit elongated rod-like shape and this was later confirmed by predictions from small angle X-ray scattering (SAXS) studies (Yang et al., 2011). Another contribution to the unusual behaviour in biochemical assays may have come from the fact that, as assessed by circular dichroism, MeCP2 demonstrates intrinsic disorder and 60% of the protein is unstructured (Adams et al., 2007). These may also be the reasons why the crystal structure of the full length MeCP2 has not been solved yet.

1.3.2 Post-translational modifications of MeCP2

One of the ways to regulate protein activity involves post-translational modifications (PTM). There is a plethora of chemical marks than can be added to the protein and they can have various effects. For example polyubiquitination tags a protein for degradation, phosphorylation can change the interaction with other proteins and SUMOylation can affect intracellular localisation. The search for PTMs in MeCP2 revealed that it can be modified on many residues. The first identified modification was phosphorylation upon neuronal membrane depolarization (Chen et al., 2003). The calcium influx-induced phosphorylation of MeCP2 caused dissociation from brain-derived neurotrophic factor (*Bdnf*) promoter III in rat cortical neurons. It was later shown that this activity-dependent brain-specific phosphorylation of MeCP2 occurs on serine 421 (Zhou et al., 2006). In neuronal cell cultures the S421 phosphorylation correlated with *Bdnf* expression activation and dendritic spine maturation.

More phosphorylation sites have been identified by mass-spectrometric analysis of immunoprecipitated MeCP2 from normal and epileptic mouse and rat brains (Tao et al., 2009). The phosphorylation on serine 80 was present constitutively, whereas S421 and newly identified S424 were phosphorylated only upon seizure induction. Disappointingly, generated mice with mutated S80, S421 and S421/S424 residues exhibited rather mild phenotypes suggesting that phosphorylation of these residues is not crucial for the function of MeCP2 (Tao et al., 2009; Cohen et al., 2011).

Bioinformatic analysis of the MeCP2 sequence revealed two PEST domains, enriched in proline, glutamate, serine and threonine (Thambirajah et al., 2009). These domains usually have lysine residues at their flanking regions and are likely to be SUMOylated and ubiquitinated, which can regulate protein activity or degradation (Rogers et al., 1986; Rechsteiner and Rogers, 1996; Melchior, 2000). Indeed, both marks have been found on MeCP2. The SUMO molecules were seen very weakly on Western blots and without mapping their positions (Miyake and Nagai, 2007). A more recent comprehensive mass spectrometry analysis of MeCP2 PTMs did not find SUMO modifications, but instead identified ten ubiquitinated lysines, some overlapping with predicted ends of the PEST domains (Gonzales et al., 2012). The authors additionally detected new phosphorylation sites in MeCP2 including those already reported. The effect of these modifications on MeCP2 function is still unknown.

1.3.3 MeCP2 binding to methylated DNA

Since its discovery, MeCP2 was considered as a protein capable of preferential binding to methylated CpG dinucleotides in DNA. There is ample evidence based on various techniques supporting this notion. Using Southwestern assays it was shown that MeCP2 binds to DNA containing a single symmetrically methylated CpG and does not bind the hemimethylated substrate or when cytosines are methylated in non-CpG context (Lewis et al., 1992; Meehan et al., 1992). Also, in these experiments MeCP2 did not interact with unmethylated DNA in the presence of a DNA competitor. Furthermore, purified MeCP2 binds preferentially to methylated DNA in electromobility shift assays (Meehan et al., 1992; Jones et al., 1998; Kaludov and Wolffe, 2000; Klose and Bird, 2004). The fragment of MeCP2 responsible for methyl-specific DNA binding has been mapped to the 78 – 162 aa region of the rat protein and was named as the methyl-CpG binding domain (MBD; Nan et al., 1993). Identification of this domain allowed the discovery of other proteins binding to methylated CpGs via MBDs (Cross et al., 1997; Hendrich and Bird, 1998; see section 1.2). The MBD domain is the only part of MeCP2 for which structural studies have been successful thus far. Nuclear magnetic resonance (NMR) analysis suggested that it has a wedge-like shape, lacking the internal symmetry,

despite the symmetry of its binding substrate, methylated-CpG (Wakefield et al., 1999). The interaction with DNA was thought to be due to van der Waals contacts between methyl groups on cytosines and a hydrophobic pocket in the MBD. However, X-ray crystallography of the MBD domain associated with methylated DNA fragment gave a similar overall structure of the domain, but revealed a surprising hydrophilic DNA-binding surface and structural water molecules facilitating the interaction between methyl-CpGs and MeCP2 (Ho et al., 2008).

The minimal size of DNA sufficient for binding of MeCP2 was estimated to be 11 base pairs (Nan et al., 1993; Nikitina et al., 2007a; Ghosh et al., 2010a). Analysis of binding preference of the MBD of MeCP2 using methyl-SELEX assay (systematic evolution of ligands by exponential enrichment) revealed that the binding to DNA was increased when methylated CpG was neighboured by a run of 4 or more adenines or thymines (Klose et al., 2005). This sequence requirement was observed only for the MBD of MeCP2 and was not necessary for the MBD domain from MBD2 protein. The increased *in vitro* binding affinity with increased occurrence of methylated CpG-$[A/T]_{\geq 4}$ motifs was also observed by sedimentation equilibrium analyses (Ghosh et al., 2010a). The X-ray crystal structure of the DNA fragment bound by the MBD of MeCP2 also supports the need of an AT-run to bend methylated DNA in order to accommodate its interaction with the 'Asx-ST' motif of MeCP2 (Ho et al., 2008). This motif is not present in methyl-CpG binding domains of other MBD proteins.

MeCP2 was also tested for interactions with chromatin and was shown to selectively bind only to methylated version of the *in vitro* chromatinised plasmid (Nan et al., 1997). The same authors suggested that MeCP2 binds to the linker DNA between nucleosomes. This notion was confirmed later by band-shift assays with reconstituted mononucleosomes and MeCP2 bound methylated nucleosome entry-exit sites better than when it was not methylated (Nikitina et al., 2007a; Ishibashi et al., 2008). DNase I footprinting analysis revealed that MeCP2 binds preferentially to methyl-CpGs exposed in the major groove of DNA on the nucleosome surface and linker DNA (Chandler et al., 1999). Another line of evidence comes from structural studies employing SAXS, where MeCP2 was able to bind to partially unpeeled nucleosomal DNA, but exhibited high preference to linker DNA (Yang et al., 2011).

Immunofluorescent staining of MeCP2 showed that it localises in the nucleus at the characteristic foci of pericentromeric heterochromatin (Lewis et al., 1992), which in mouse is known to be rich in methylated microsatellite sequences (Miller et al., 1974; Hörz et al., 1981). In mouse cells with 95% lower levels of DNA methylation due to a mutation in DNA methyltransferase gene, MeCP2 was not localised in these foci (Nan et al., 1996). Additionally, treatment of cells with the demethylating agent, 5-azadeoxycytidine, resulting in globally reduced DNA methylation, caused depletion of MeCP2 from heterochromatin (Ghosh et al., 2010a). In cells of species lacking the pericentromeric heterochromatin (rat, human, monkey) MeCP2 coated whole chromosomes (Lewis et al., 1992; Nan et al., 1997) in agreement with the fact that the majority of the mammalian genome is methylated (Illingworth and Bird, 2009).

There is also ample evidence for MeCP2 binding preferentially to methylated DNA *in vivo*, coming from ChIP experiments on imprinted genes or differentially methylated chromatin (El-Osta and Wolffe, 2001; Gregory et al., 2001; Lorincz et al., 2001; Nguyen et al., 2001; El-Osta et al., 2002; Rietveld et al., 2002; Matarazzo et al., 2007; Muotri et al., 2010). Furthermore, a detailed study from our laboratory showed MeCP2 binding across the *Xist* locus and peaking over its methylated CpG island in male mouse brain (Skene et al., 2010). In females, who have two copies of the gene, one is methylated, whereas the other allele remains unmethylated to facilitate X chromosome inactivation (Beard et al., 1995; McDonald et al., 1998). Binding of MeCP2 over the *Xist* CpG island in female mouse brains was reduced 2-fold when compared to the male binding profile, while in the flanking regions the profiles were indistinguishable. Bisulphite sequencing of female input DNA confirmed methylation of 50% of clones, but the analysis of material immunoprecipitated with MeCP2 revealed that almost 90% of MeCP2-bound chromatin was methylated (Skene et al., 2010).

1.3.4 MeCP2 binding to unmethylated DNA

Contrary to the publications presented above, some researchers postulate that MeCP2 can also bind strongly to unmethylated DNA. The evidence for this comes

mostly from *in vitro* studies. The interaction with nucleosomal arrays regardless of their cytosine methylation status was ascribed to the regions of MeCP2 outside of the MBD domain, mainly in the C-terminal part of the protein (Georgel et al., 2003; Nikitina et al., 2007b; Ghosh et al., 2010b). Considering the strong basic character of MeCP2, the interaction with any negatively charged DNA is not surprising, especially given that in these studies no DNA competitor (Georgel et al., 2003), or only 2 – 3 - fold excess over the tested probe (Nikitina et al., 2007b; Ghosh et al., 2010b) were used. In another *in vitro* study MeCP2 was shown to interact with unmethylated 4-way junction DNA via its MBD domain, even in the presence of a large excess of competitor (Galvão and Thomas, 2005). Although the nature of this binding has not been determined, it was suggested that MeCP2 is able to recognise the cruciform structure of the junction DNA and has higher affinity to it over DNA in linear form.

There is, however, a remaining question of how relevant these types of interactions are *in vivo*. The binding of MeCP2 to unmethylated active genes has been reported in human neuronal cell line SH-SY5Y (Yasui et al., 2007). Notably, the antibody used in the ChIP-microarray analyses in this study showed low enrichment of MeCP2-bound material in wild-type over *Mecp2*-null brains, thus questioning the validity of authors' conclusions. Two other studies using ChIP at specific chromatin regions revealed that MeCP2 can bind to the unmethylated maternal *H19* imprinting control region (Kernohan et al., 2010) and to the non-methylated CpG island promoter of *Slc6a2* gene (Harikrishnan et al., 2010). However in both cases MeCP2 was associated with large proteins: ATRX and cohesin (Kernohan et al., 2010), and the Smarca2 complex (Harikrishnan et al., 2010), which could influence the binding specificity of MeCP2.

The rapid development of high-throughput sequencing methods as well as chromatin conformation capture strategies should help in gaining a better understanding of MeCP2 binding specificity and distribution *in vivo*. In fact, quite recently MeCP2 was immunoprecipitated from mouse brains and the fragments of chromatin bound by it were subjected to the next generation sequencing (Skene et al., 2010). The depth of sequencing in this study allowed only low-resolution mapping of MeCP2 distribution; nevertheless it indicated that the binding profile of MeCP2

tracks the profile of methylated cytosine within the genome. Interestingly, even in chromatin regions known to be unmethylated (such as CpG islands of active genes) the binding of MeCP2 was always higher than the non-specific antibody binding in the control MeCP2 knock-out brains (Skene et al., 2010). These residual interactions with non-methylated DNA may be explained by the high abundance of MeCP2 in neurons causing it to bind non-specifically. Additionally, a search of the obtained ChIP-seq data for presence or absence of the CpG-$[A/T]_{\geq 4}$ motif in DNA fragments bound by MeCP2 revealed no preference for such sequences (Skene, unpublished). It is possible, however, that the limitations of ChIP-seq technique masked the identification of specific MeCP2 binding sites. The immunoprecipitated fragments were ~500 bp long and contained more than one CpG per fragment, therefore even if MeCP2 was binding only to CpG-$[A/T]_{\geq 4}$ motifs, the sequencing of the whole fragment would not reveal this binding sequence requirement. Use of ChIP-exo technique in the future could overcome this potential problem (Rhee and Pugh, 2011). Further experiments aimed at unravelling the genome-wide binding specificity of MeCP2 are ongoing in ours and other laboratories.

1.3.5 MeCP2 as a transcriptional repressor

Methylation of DNA at CpG dinucleotides is considered as a mark of a transcriptionally silent chromatin state (Stein et al., 1982). The discovery of methyl-CpG binding proteins led to an outburst of research suggesting their role in the transcriptional repression (Kass et al., 1997). MeCP2 was not different and although initial studies did not demonstrate its methyl-specific silencing character (Meehan et al., 1992), an improved set of experiments later revealed that MeCP2 acts as a transcriptional repressor (Nan et al., 1997). The minimal fragment of MeCP2 able to repress transcription of a reporter gene has been mapped to the 207-310 aa region of the rat protein and named as the transcriptional repression domain (TRD; Nan et al., 1997). MeCP2 was able to repress transcription not only when binding to the promoter of reporter gene, but also from a distance up to 2 kb from the transcriptional start site (Nan et al., 1997; Kaludov and Wolffe, 2000).

The mechanism of function of MeCP2 as a repressor of transcription was proposed based on the discovery that MeCP2 interacts with co-repressor complex Sin3A in HeLa cells and rat brains (Nan et al., 1998) as well as in *Xenopus* oocytes (Jones et al., 1998). The Sin3A complex contains histone deacetylases and in line with this fact were observations in both studies that the repression of reporter genes was relieved by treatment with trichostatin A (TSA), an HDAC inhibitor (Yoshida et al., 2007). Notably, the region of MeCP2 mapped to interact with Sin3A was not entirely overlapping with the functionally mapped transcriptional repression domain and the interaction is relatively weak (M. Lyst, personal communication) and not stable in biochemical assays (Klose and Bird, 2004). There were also other mechanisms proposed, based on the detection of binding of MeCP2 to various repressor proteins, such as Ski, NCoR (Kokura et al., 2001); SMRT (Stancheva et al., 2003); CoREST, SUV39H1 (Lunyak et al., 2002); Brahma (Smarca2) (Harikrishnan et al., 2005, 2010) and YY1 (Forlani et al., 2010). Additionally, repression of transcription mediated by MeCP2 binding the transcription factor IIB (Kaludov and Wolffe, 2000) and in an HDAC-independent manner (Yu et al., 2000) have been observed, however both studies used artificial *in vitro* systems and the relevance of these discoveries *in vivo* is not known. Surprisingly, none of the aforementioned interactions of MeCP2 has been followed in detail in subsequent studies. The validity of the MeCP2 binding to Brahma has been questioned as it could not be confirmed by other laboratories (Harikrishnan et al., 2006; Hu et al., 2006). Therefore, up to now the mechanism of MeCP2-mediated transcriptional repression was generally believed to be due to its interaction with the Sin3A co-repressor (Klose and Bird, 2006; Chahrour et al., 2008).

The functional link between two domains of MeCP2 has been demonstrated by means of methylated reporter assays. In these experiments transcription of a reporter gene with methylated CpGs is compared to the transcription of the same gene in the unmethylated state. As mammalian differentiated cells express various methyl-CpG binding proteins the transcription of luciferase reporter transfected into mouse tail fibroblasts was repressed by its methylation. Genetic knock-out of two genes coding MBD-containing proteins, MeCP2 and MBD2, resulted in an almost 30-fold increase of normalised luciferase activity (Guy et al., 2001). Co-transfection

of these cells with a construct expressing MeCP2 reduced reporter gene activity to the normalised level of wild-type cells, suggesting that MeCP2 binds to the methylated reporter and represses its transcription (Guy et al., 2001). Similar experiments performed in HeLa cells confirmed that the repression of methylated reporter gene by MeCP2 is mediated by its TRD domain (Drewell et al., 2002). The repression was alleviated by TSA treatment, again suggesting the involvement of histone deacetylases in the process.

1.3.6 Genes regulated by MeCP2

Availability of *Mecp2* knock-out mice and the rapid development of molecular tools to globally investigate transcriptomes have allowed the search for putative MeCP2 target genes. Microarray studies comparing transcripts in wild-type mouse brains and in *Mecp2*-null mouse found only a handful of subtle differences (Tudor et al., 2002). Subsequent extensive searches employing various detection platforms have managed to identify several genes mis-regulated in *Mecp2*-null mice, but there was no consensus between the studies (Nuber et al., 2005; Kriaucionis et al., 2006; Deng et al., 2007; Jordan et al., 2007; Chahrour et al., 2008; Urdinguio et al., 2008; Ben-Shachar et al., 2009). Moreover, one study compared the transcriptomes of two different mouse models of RTT at different timepoints and the overlap was minimal (Jordan et al., 2007). Similar searches were performed in human samples from patients suffering from Rett syndrome. Analysis of gene expression in post-mortem brains (Colantuoni et al., 2001), lymphocytes (Delgado et al., 2006), lymphoblastoid cells (Traynor et al., 2002; Ballestar et al., 2005) and fibroblasts (Traynor et al., 2002; Nectoux et al., 2010) again did not provide any obvious or overlapping MeCP2 target genes.

Despite numerous trials using various approaches the data obtained was surprisingly inconclusive and failed to determine a reliable set of genes regulated by MeCP2. The relatively small changes in expression levels of tested genes between wild-type and KO brains may be due to the fact that brain is a mixture of different cell types and even neurons of the same type are not activated simultaneously. That could mask the effect of MeCP2 deletion if its role is to regulate transcription of

genes only upon activation of neurons. The candidate gene approach focusing on genes involved in neuronal function was somewhat more successful. One gene in particular, brain-derived neurotrophic factor (*Bdnf*), has been reported by several groups as mis-regulated in MeCP2 deficient cells (Chen et al., 2003; Martinowich et al., 2003; Chang et al., 2006; Chahrour et al., 2008; Larimore et al., 2009). BDNF is a protein with complex functions, involved in many processes in the brain such as neuronal development, survival and plasticity (reviewed in: Cohen-Cory et al., 2010). Transcription of *Bdnf* gene is regulated by 8 alternative promoters and shows activity dependence and cell-type specificity (Aid et al., 2007). MeCP2 has been found to bind to the *Bdnf* promoter at exon III in rat cortical neurons (Chen et al., 2003), and to a corresponding promoter in the mouse, at exon IV (Martinowich et al., 2003) and repress the basal transcription. Upon neuronal membrane depolarisation the MeCP2 dissociates from the promoter, probably due to phosphorylation at S421 (Chen et al., 2003; Zhou et al., 2006).

Due to the high abundance of MeCP2 and its binding motif, methyl-CpG, it is distributed broadly across the neuronal chromatin (Lewis et al., 1992; Skene et al., 2010; Cohen et al., 2011). This genome-wide binding profile and the fact that MeCP2 can act as a transcriptional repressor in *in vitro* reporter gene assays suggest that it may work globally rather than on a specific subset of genes. The idea of MeCP2 serving as a general 'dampener' of spurious transcription was supported by detection of increased transcription from the repetitive elements in nuclei from *Mecp2*–null brains (Skene et al., 2010). Additionally, the global levels of histone H3 acetylation were increased in the sorted neuronal nuclei of *Mecp2* KO mouse (Skene et al., 2010) and in various brain structures of $Mecp2^{308/y}$ mutant mouse (Shahbazian et al., 2002a) implicating its global function through histone deacetylases.

The functional consequence of such a global transcriptional noise from genomic regions not coding proteins is unknown. The intronic, intergenic and repetitive sequences quite often contain retrotransposons or regulatory elements such as microRNAs or long non-coding RNAs, which when expressed may affect other proteins and various cellular processes. The role of MeCP2 in regulation of expression of microRNAs was tested, but yet again a consensus has not been found between the different studies (Szulwach et al., 2010; Urdinguio et al., 2010;

Wu et al., 2010). Future experiments using unbiased high-throughput RNA sequencing should help in identifying the effect of MeCP2 on global and local transcription.

1.3.7 MeCP2 as a transcriptional activator

Surprisingly, it has been noticed that in the hypothalamus of *Mecp2* knock-out mice more genes are mildly downregulated than upregulated, suggesting that MeCP2 may act like a transcriptional activator rather than repressor (Chahrour et al., 2008). This study was further supported by analysis of gene expression in hypothalamus of mice over-expressing MeCP2 two-fold ($MeCP2^{Tg1}$; Collins et al., 2004). In these animals the majority of downregulated genes in *Mecp2*-null mice showed reciprocal changes of expression and were upregulated and vice versa (Chahrour et al., 2008). Another study from the same laboratory focused on a different brain region, the cerebellum and the findings were comparable and a subset of identified genes activated by MeCP2 overlapped (Ben-Shachar et al., 2009). The mechanism of this transcriptional activation by MeCP2 was proposed based on the putative interaction of MeCP2 with CREB1 (cyclic-AMP responsive element binding protein 1), though it was only weakly detected by mass spectrometry (Chahrour et al., 2008). Bisulphite sequencing and limited ChIP-PCR analysis revealed that in wild-type brain MeCP2 was bound to the unmethylated promoters of the genes upregulated upon its deletion. ChIP-on-chip experiments in the human neuronal cell line SH-SY5Y revealed that MeCP2 binds to unmethylated promoters of active genes (Yasui et al., 2007). Surprisingly, the MeCP2 binding profile frequently overlapped with distribution of RNA polymerase II.

The whole effect of decreased transcription of some genes upon deletion of *Mecp2* may well be indirect. It is possible to imagine wild-type situation when MeCP2 is engaging the transcriptional repression machinery throughout the bulk methylated genome to repress spurious transcription. In consequence, the genes coding the functional proteins are not repressed and are transcribed at their physiological level. In the case of depletion of MeCP2 the co-repressors normally bound to methylated DNA by MeCP2 are released and may then act on the promoters of functional genes, resulting in their decreased transcription. This would

be detected by microarrays in *Mecp2* KO brains suggesting the role of MeCP2 as a transcriptional activator. This speculative hypothesis remains to be tested.

1.3.8 Other functions of MeCP2

Assembly of higher-order chromatin structures

Apart from the best studied MeCP2 function as a transcriptional regulator it has been also implicated in many other processes. The binding to methylated and unmethylated DNA through different domains has suggested a possible role in chromatin organisation. Indeed, electron microscopy of artificial chromosomes revealed that MeCP2 is able to condense chromatin to higher-order structures (Georgel et al., 2003; Nikitina et al., 2007a, 2007b). The DNA loops created by single MeCP2 molecules bridging distant fragments of naked DNA were also observed *in vitro* using atomic force microscopy (Ghosh et al., 2010a). The only study *in vivo* presenting MeCP2-mediated forming of chromatin loops employed the chromatin conformation capture method. Horike and co-workers showed that MeCP2 facilitated forming of an 11 kb-long loop at the *Dlx5-Dlx6* locus and in consequence silencing the imprinted allele (Horike et al., 2005). The findings of these authors, however, were later challenged (Schüle et al., 2007).

Replacement of histone H1

MeCP2 has also been shown to be able to displace histone H1 from *in vitro* chromatinised methylated plasmid (Nan et al., 1997). Interestingly, DNase I footprinting analysis of H1 and MeCP2 binding to an *in vitro* reconstituted nucleosome revealed that these two proteins have distinct binding pattern (Chandler et al., 1999). However, this observation is based only on one DNA sequence tested (5S rRNA gene) and may not be true globally.

Analysis of changes in fluorescent anisotropy of labelled MeCP2 or H1 upon incubation with unlabelled protein showed that MeCP2 is more potent in displacing H1 than vice versa (Ghosh et al., 2010a). These experiments suggested that both proteins share the same binding sites and can not co-exist on the same nucleosomal linker DNA. These observations were confirmed *in vivo* by recording changes in

fluorescence recovery after photobleaching (FRAP) kinetics of H1-GFP when challenged by microinjection of MeCP2 into the nucleus and vice versa (Ghosh et al., 2010a).

A potential link between the function of MeCP2 and H1 comes from the fact that although histone H1 is present at approximate stoichiometry of one molecule per nucleosome in most tissues (Woodcock, 2006), its abundance in neurons is reduced two-fold (Pearson et al., 1984). The number of MeCP2 molecules in neurons was estimated at approximately one per two nucleosome (Skene et al., 2010) as if matching the missing number of H1 molecules. Interestingly, the expression of histone H1 in neurons of *Mecp2* knock-out mice was doubled when compared to the wild-type levels, probably compensating the lack of MeCP2 on the nucleosomal linkers (Skene et al., 2010).

All these findings suggest the possibility that, at least partially, MeCP2 and H1 have similar functions in the maintaining the structure of chromatin.

Prevention of retrotransposition

Two laboratories reported a relationship between MeCP2 and long interspersed nuclear elements-1 (LINE-1), retrotranspons abundantly distributed in the genome (Yu et al., 2001; Muotri et al., 2010). MeCP2, but not MBD1 or MBD2, was able to repress transcription of a reporter gene with the methylated promoter of the L1 retrotransposon. Moreover, the retrotransposition assay revealed that expression of MeCP2 in HeLa cells reduced the frequency of retrotransposition events of the methylated construct (Yu et al., 2001). A well designed set of experiments performed in Gage laboratory confirmed these results and demonstrated increased retrotransposition of LINE-1 elements in MeCP2-deficient cultured neurons and mouse brains (Muotri et al., 2010). Additionally, the authors tested the effect of MeCP2 on LINE-1 retrotransposition in neuronal progenitor cells derived from induced pluripotent stem cells generated from fibroblasts of RTT patients. Again, more retrotransposition events were detected in MeCP2 deficient cells and it was also found to be true in post-mortem brain samples, but not in matched cardiac tissue from the same patients. The role of LINE-1 retrotransposition on the scale of the whole organism and their relevance to the Rett syndrome are not known. These

studies, however, add another layer of complexity towards understanding the function of MeCP2.

Regulation of alternative pre-mRNA splicing

Another potential role of MeCP2 was proposed upon discovery of its interaction with Y-box binding protein 1 (YB-1; Young et al., 2005). YB-1 is a protein involved in many cellular processes including regulation of transcription, translation, DNA repair and RNA splicing (Kohno et al., 2003). The MeCP2-YB-1 interaction was shown to be independent of binding to methylated DNA and was mapped to the TRD domain and C-terminal end of MeCP2. Interestingly it was abolished by RNAse treatment, suggesting the involvement of RNA in the binding. MeCP2 has been also reported to interact with RNA before (Jeffery and Nakielny, 2004). Young and co-authors suggested that MeCP2 together with YB-1 may regulate alternative splicing and supported their hypothesis using reporter minigenes. Additionally, a comparison of mRNAs from mouse brains expressing wild-type and a truncated form of MeCP2 (*Mecp2*$^{308/y}$) on a custom-made splicing microarray revealed a number of genes spliced differently. The fact that another mutated form of MeCP2 (R106W) that is unable to bind to methylated DNA still interacts with RNA and can regulate its splicing (Jeffery and Nakielny, 2004; Young et al., 2005) argues against the relevance of these discoveries for a global picture of MeCP2 function. Two other proteins potentially involved in alternative splicing have been identified to interact with MeCP2, formin-binding protein 11 (FBP11), a splicing factor containing the WW domain (Bedford et al., 1997) and huntingtin-interacting protein C (HYPC; Buschdorf and Strätling, 2004). It allowed a discovery of a WW-domain binding region on MeCP2 at its C-terminus, responsible for these interactions and possibly binding other splicing factors (Buschdorf and Strätling, 2004). However, the significance of these interactions *in vivo* is not known.

In summary, MeCP2 appears as a complex protein interacting with DNA and various other proteins. There are numerous potential functions ascribed to it but it is not known which of them are the most relevant *in vivo*. The insights for deciphering the key mechanism of action of MeCP2 should come from studying its association with Rett syndrome.

1.4 Rett syndrome

1.4.1 Rett syndrome in humans

Rett syndrome (RTT) was described for the first time by Austrian paediatrician and neurologist Andreas Rett in 1966, but was not published in English until 1983 (Hagberg et al., 1983). This neurological disorder occurs approximately 1 : 12 500 births and affects mostly girls, who develop normally for 6-18 months when they start exhibiting a range of symptoms and lose acquired skills, such as speech and walking (Cheval et al., 2012). The classical Rett patients suffer from stereotypic hand movements, seizures, gait difficulties, ataxia, breathing abnormalities, mental retardation and autism-like social interactions (Hagberg et al., 1983). It was reported that the syndrome is an X-linked disease and it was mapped to Xq28 region (Sirianni et al., 1998; Xiang et al., 1998). The candidate gene approach allowed the identification of a gene mutated in the Rett syndrome patients, which turned out to be *MECP2* (Amir et al., 1999). More than 90% of sporadic cases of classical RTT in females are as a result of mutations in the *MECP2* gene (Bienvenu and Chelly, 2006). All of the genetic alterations identified in RTT patients were compiled in RettBASE database (Christodoulou et al., 2003). The general understanding is that disease-causing mutations can occur throughout the whole length of protein (Kriaucionis and Bird, 2003; Adkins and Georgel, 2011; Chao and Zoghbi, 2012). The search for a correlation between *MECP2* mutation and Rett syndrome symptoms severity failed to find any clear links (Temudo et al., 2011) suggesting that all regions of MeCP2 are important for its function. Such a global view of all types of mutations does not provide specific information about MeCP2, because the nonsense and frameshift mutations disrupt large part of the protein having effect on its various aspects, such as folding, stability, localisation or multiple interactions. However, focusing only on missense mutations allows precise identification of the functional sites on MeCP2 that are key to Rett syndrome, because unlike other mutations, they surgically affect just one residue. Moreover, in some cases of RTT the parents of patients were not tested for mutations in their *MECP2* gene and therefore the reported mutations could simply be polymorphisms not related to a disease. Careful analysis of the database regarding only the *de novo*

missense mutations revealed two clusters of mutational hotspots (Fig. 1.4). One of the large regions commonly mutated in RTT is restricted to the MBD domain and plausibly affects MeCP2 binding to methylated DNA. Both deletion and missense mutations in this domain cause disruption of binding to DNA as assessed by Southwestern (Ballestar et al., 2000; Yusufzai and Wolffe, 2000) and band-shift assays (Ballestar et al., 2000; Nikitina et al., 2007a, 2007b; Ho et al., 2008). The characteristic localisation of MeCP2 in heterochromatic foci was more diffuse when the protein was mutated in the MBD, as observed by immunofluorescent microscopy in fixed cells (Kudo et al., 2001, 2003; Schmiedeberg et al., 2009). Interestingly, live-cell imaging of cells expressing mutated versions of MeCP2 showed their wild-type-like localisation (Schmiedeberg et al., 2009) and it was suggested that the differences lie rather in their binding dynamics (Kumar et al., 2008; Schmiedeberg, unpublished).

The other cluster of RTT-causing missense mutations is localised at the C-terminus of the TRD domain and we hypothesise that these mutations could affect binding of repressor complexes responsible for MeCP2-mediated transcriptional repression.

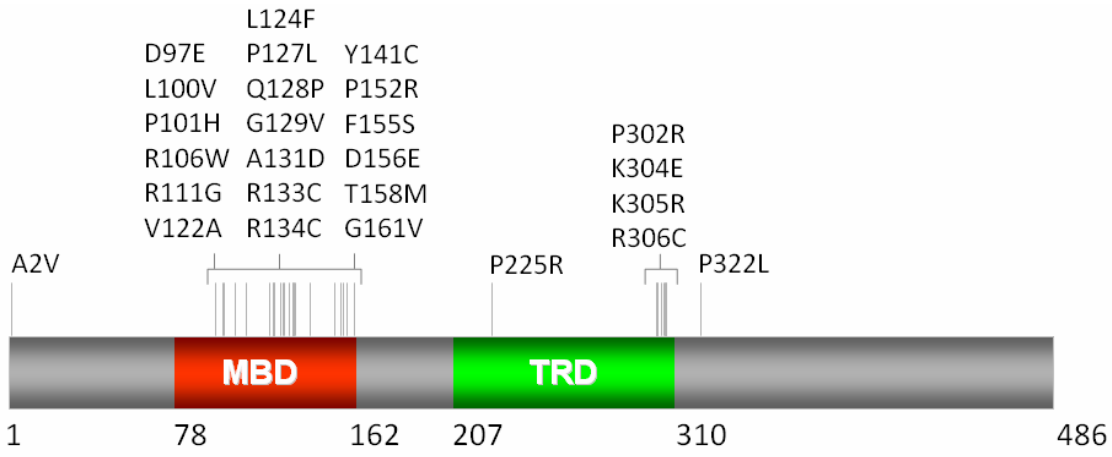


Figure 1.4. RTT missense mutations cluster in the functional domains of MeCP2.
De novo missense mutations causing Rett syndrome were collected from RettBASE (Christodoulou et al., 2003) and depicted on the human MeCP2 isoform e2 sequence. MBD - methyl-CpG binding domain, TRD - transcriptional repression domain.

1.4.2 Mouse models of Rett syndrome

The first knock-out the *Mecp2* gene in mice was performed before the link between MeCP2 protein and the Rett syndrome was known. The substitution of *Mecp2* with *lacZ* gene in embryonic stem (ES) cells did not affect their morphology and differentiation potential. However, creation of *Mecp2*-null mouse from these cells was not successful due to embryonic lethality (Tate et al., 1996) suggesting its critical role in development. After the discovery of the genetic cause of RTT (Amir et al., 1999), two groups decided independently to delete *Mecp2* using Cre-*loxP* conditional knock-out approach (Chen et al., 2001b; Guy et al., 2001). The resulting mice recapitulated most of the human Rett syndrome phenotypes as the apparently normal development of mouse was followed by a delayed onset of neurological symptoms including tremors, uncoordinated gait, abnormal breathing, and reduced brain size. Null animals started exhibiting symptoms between 3-8 weeks and died on average on day 54 after birth (Guy et al., 2001). The onset of symptoms in female heterozygous mice was delayed to 3-9 months and they were fertile and exhibited normal lifespans. Both homo- and heterozygotes for MeCP2 deletion are valid mouse models of Rett syndrome, however most subsequent studies were performed on male null mice as the phenotype appears faster and more uniformly.

Conditional deletion of *Mecp2* in whole brain by *Nestin*-Cre gave rise to a phenotype comparable with the full body knock-out (Chen et al., 2001b; Guy et al., 2001). However, *CamKII*-Cre-mediated *Mecp2* deletion in post-mitotic neurons of forebrain, hippocampus and brainstem, but not in cerebellum and glial cells resulted in delayed and less severe phenotype (Chen et al., 2001b). Other region- and cell-type specific *Mecp2* deletions in brain recapitulated some of the RTT phenotypes to various extents (Reviewed in: Gadalla et al., 2011 and Guy et al., 2011). *MECP2* gene knocked in the *Tau* locus of *Mecp2* null mice was able to rescue the RTT phenotype (Luikenhuis et al., 2004) suggesting the role of MeCP2 in the post-mitotic neurons. Interestingly, the function of MeCP2 in glial cells have also been recently reported (Lioy et al., 2011; Derecki et al., 2012).

The remaining question with tremendous consequences for potential RTT therapies was whether MeCP2 deletions affect neuronal development, cause their irreversible degeneration or just impair their function. The answer came from very neat experiments when RTT mice were produced by inserting the floxed STOP-Neo cassette before exon 3 of *Mecp2* and crossing it with a mouse inducibly expressing Cre recombinase upon treatment with tamoxifen (Guy et al., 2007). The mice were allowed to develop RTT symptoms and subsequent tamoxifen injections caused Cre to be expressed and excise the STOP cassette allowing full-length MeCP2 to be produced. Spectacular reversal of the advanced neuronal phenotypes has been observed (Guy et al., 2007). This suggested that MeCP2 is crucial for the proper neuronal function, but lack of it does not impair them irreversibly. This discovery has tremendous importance for the search of potential treatments for Rett syndrome.

The precise regulation of MeCP2 levels in cells is critical as too much and too little of the protein is detrimental. Humans with duplication of a locus containing *MECP2* gene exhibit neurological symptoms, severe mental retardation, limited or absent speech, spasticity, facial hypotonia and locomotor problems (Van Esch et al., 2005). Transgenic mice expressing twice as much MeCP2 as wild-type levels ($MeCP2^{Tg1}$) also demonstrate a spectrum of neurological and behavioural problems and premature death (Collins et al., 2004). Moreover, in mice with three-fold MeCP2 expression ($MeCP2^{Tg3}$) the phenotypic severity is increased. On the other hand, reducing the level of MeCP2 by 50% results in mild, but statistically significant differences in mice behaviour and performance (Samaco et al., 2008).

Because of the necessity of keeping the expression of MeCP2 at wild-type levels, the potential gene therapy aimed at replacing the faulty *MECP2* gene in Rett syndrome patients may be difficult (Reviewed in: Gadalla et al., 2011). The fact that in RTT females approximately half of the neurons are expressing normal levels of MeCP2 presents a real hurdle in the design and/or proper targeting the gene therapy only to the defective cells. Therefore, despite the initial trials in mice serving rather as a proof of principle, the prospective therapy based on gene delivery is still far away from clinics.

A very recent study gives some hope to prevent Rett syndrome. Experiments in *Mecp2* null mice showed that a bone marrow transplant from wild-type animal can prevent the occurrence of RTT-like pathology (Derecki et al., 2012). The proposed mechanism relies on migration of transplanted microglial cells in blood, crossing the blood-brain barrier and serving as phagocytes removing cell debris in the brain. The drawback of this approach is that it was only successful in arresting the symptoms occurrence in pre-symptomatic mice. To apply it to humans, every newborn child should be screened for RTT mutations and given a bone marrow transplant before it can be diagnosed with the syndrome based on the symptoms. It sounds unrealistic at the moment, but rapidly dropping prices and increasing speed of genome sequencing should make this kind of approach feasible in the future.

Aims of the thesis

Despite more than 20 years of research on MeCP2, its molecular function and the role it plays in Rett syndrome are still not fully known. Many putative mechanisms of action have been ascribed to MeCP2, however it is important to distinguish between the functions which may be true when tested in *in vitro* assays, and the functions found and proved *in vivo*. In my opinion the most interesting and relevant are the MeCP2 activities which are disrupted in Rett syndrome. Therefore, in this study we decided to re-address the issue of deciphering the function of MeCP2, focussing on its association with RTT.

The aims of this thesis were:

(1) To search for interacting partners of MeCP2 and identify new proteins
(2) To map the confirmed interactions onto the MeCP2 sequence
(3) To investigate the functionality of these interactions with regard to Rett syndrome causing mutations

Chapter 2: Materials and methods

2.1 Materials

All standard chemicals and reagents used in this study were provided by Sigma-Aldrich and all tissue culture media and chemicals were supplied from Gibco, unless stated otherwise.

2.1.1 Standard buffers

DNA sequencing buffer (2.5×):
20 mM Tris-HCl, 5 mM $MgCl_2$, pH 8.0

Orange G DNA loading buffer (5×):
0.198% orange G, 12% Ficoll, 120 mM EDTA, 4.2% SDS, pH 8.0

PBS (phosphate buffered saline):
137 mM NaCl, 2.7 mM KCl, 2 mM KH_2PO_4, 10 mM $Na_2HPO_4 \cdot 2H_2O$, pH 7.4

SDS-PAGE sample buffer (2×):
125 mM Tris-HCl (pH 6.8), 4% SDS, 20% glycerol, 300 mM β-mercaptoethanol, 0.2% bromophenol blue

SDS-PAGE running buffer:
25 mM Tris, 250 mM glycine, 0.1% SDS

Transfer buffer:
25 mM Tris, 192 mM glycine

TAE (Tris-acetate EDTA):
40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0

TBS (Tris-buffered saline):
50 mM Tris-HCl, 150 mM NaCl, pH 8.0

TBS-T (TBS-Tween):
TBS + 0.05% Tween 20 (Bio-rad)

2.1.2 Antibodies

All primary antibodies used in this study are listed in the Table 2.1.2A and the secondary antibodies in the Table 2.1.2B.

Table 2.1.2A Primary antibodies used in this study. The symbols for applications are as follows: WB – Western blot, IHC – immunohistochemistry, IP – immunoprecipitation.

Antibody	Species	Source, Catalogue number, Stock concentration	Application (dilution)
anti-CamKIIα	mouse	Abcam, ab22609, 1 mg/ml	WB (1:1000)
anti-CK2α	rabbit	Abcam, ab10466, 1 mg/ml	WB (1:1000)
anti-FLAG	mouse	Sigma, F3165, 2 mg/ml	WB (1:2000)
anti-GAL4-DBD	mouse	Santa Cruz, Sc-510, 0.2mg/ml	WB (1:500-1:1000)
anti-GAPDH	mouse	Abcam, ab9484, 1 mg/ml	WB (1:1000)
anti-GFAP	mouse	Sigma, G3893, 2 mg/ml	IHC (1:200)
anti-H2B	rabbit	Abcam, ab1790, 0.5 mg/ml	WB (1:2000)
anti-H3	rabbit	Abcam, ab1791, 1 mg/ml	WB (1:5000)
anti-H3 acetyl	rabbit	Millipore, 06-599, 1 mg/ml	WB (1:1000)
anti-HDAC3	rabbit	Abcam, ab7030, 1 mg/ml	IHC (1:100)
anti-KPNA3	goat	Abcam, ab6038, 0.5 mg/ml	WB (1:1000) IHC (1:100)
anti-KPNA4	rabbit	Abcam, ab81451, 1 mg/ml	WB (1:1000) IHC (1:100)
anti-KPNA4	rabbit	Bethyl, A301-627A, 0.2 mg/ml	WB (1:1000)
anti-KPNA4	rabbit	Bethyl, IHC-00321, 0.2 mg/ml	IHC (1:100)
anti-MAP1B	mouse	Abcam, ab3095, 0.2 mg/ml	WB (1:1000)
anti-MAP2	rabbit	Millipore, AB5622, 1 mg/ml	IHC (1:200)
anti-MeCP2	mouse	Sigma, M6818, 2 mg/ml	WB (1:1000)
anti-MeCP2	rabbit	Millipore, 07-013, 1 mg/ml	IHC (1:100)
anti-NeuN	mouse	Millipore, MAB377, 1 mg/ml	WB (1:1000) IHC (1:400)
anti-Tuj1	mouse	Covance, MMS-435P, 1mg/ml	IHC (1:400)
anti-WBP11	rabbit	Abcam, ab85563, 1 mg/ml	WB (1:1000)
GFP-Trap_A	llama	Chromotek, gta-20	IP (1:33)
anti-FLAG M2 affinity agarose	mouse	Sigma, A2220	IP (1:200)

Table 2.1.2B Secondary antibodies used in this study. The abbreviations used here are as follows: WB – Western blot, ECL – enhanced chemiluminescence, HRP – horseradish peroxidase, IHC – immunohistochemistry.

Antibody	Source, Catalogue number	Application (dilution)
anti-mouse-HRP	GE Healthcare, NA931V	WB-ECL (1:8 000)
anti-rabbit-HRP	GE Healthcare, NA934V	WB-ECL (1:8 000)
anti-goat-HRP	Abcam, ab6885	WB-ECL (1:5 000)
anti-mouse IRDye 800CW	Odyssey, 926-32212	WB-Licor (1:10 000)
anti-rabbit IRDye 700LT	Odyssey, 926-32223	WB-Licor (1:10 000)
anti-rabbit IRDye 800CW	Odyssey, 926-32213	WB-Licor (1:10 000)
anti-goat IRDye 800CW	Odyssey, 926-32214	WB-Licor (1:10 000)
anti-mouse-594 (goat)	Molecular probes, A11032	IHC (1:500)
anti-mouse-633 (goat)	Molecular probes, A21071	IHC (1:500)
anti-rabbit-594 (donkey)	Molecular probes, A21207	IHC (1:500)
anti-goat-594 (donkey)	Molecular probes, A11058	IHC (1:500)

2.1.3 Mice

All mice used in this study were bred and maintained at the University of Edinburgh animal facilities under standard conditions and all procedures were carried out by staff licensed by UK Home Office and according to the Animals and Scientific Procedures Act 1986. The genotypes of mice used in different experiments are listed in the Table 2.1.3. All animals were 6-9 weeks old and wherever possible the matching littermates were used.

Table 2.1.3 Mice used in this study. The mouse genotypes used in each type of experiment.

Type of experiment	Mouse genotypes	Genetic background
Immunoprecipitations	♂ wild-type	C57BL/6
	♂ *Mecp2-GFP*	C57BL/6 segregating
	♂ *H2B-GFP*	(Jiang et al., 2008)
Immunohistochemistry	♀ *Mecp2-GFP*$^{+/-}$	mixed (outbred)
Analysis of protein and mRNA levels	♂ *Mecp2* $^{-/y}$	C57BL/6
	♂ wild-type	C57BL/6

2.1.4 Cell lines

The wild-type mouse ES cell line used was E14 TG2α. The *Mecp2*-GFP, *Mecp2*-GFP [T158M] and *Mecp2*-GFP [R306C] lines were generated as described elsewhere (Thomson et al., 2010). The *Mecp2*-null ES cells were produced from cells described in (Guy et al., 2001), by deletion of floxed segment by *in vitro* Cre transfection (performed by J. Guy). Other cell lines used were HeLa and NIH-3T3, cultured as described below.

2.1.5 PCR primers

All primers have been synthesised by Sigma-Aldrich, resuspended in H_2O to a concentration of 100 μM and stored at -20 °C as a stock solution. The sequences of primers are listed at each method where they were used.

2.2 Methods

The methods are listed in order of their appearance in the results chapters 3-6.

2.2.1 Isolation of nuclei from mouse brains

Three brains from 6-9 weeks old male *Mecp2-GFP* knock-in mice or age-matching wild-type (C57BL/6 strain) or *CamKII-H2B-GFP* mice were collected, snap-frozen in liquid nitrogen and stored at -80 °C. Brains thawed on ice were homogenised in 9 ml of ice-cold buffer A (10 mM HEPES pH 7.9, 25 mM KCl, 1 mM EDTA, 0.5 mM EGTA, 2 M sucrose, 10% glycerol, 0.15 mM spermine, 0.5 mM spermidine and complete protease inhibitors [Roche]) in a 15 ml glass dounce using Potter S motorised homogeniser (Braun) at 1100 rpm for seven strokes. The homogenates were layered onto 3 ml cushion of buffer A in $^{9}/_{16}$" × 3¾" polyallomer tubes (Beckman) and centrifuged at 24 000 rpm at 4 °C for 40 min on SW-40Ti rotor in Beckman Coulter XL-100 ultracentrifuge. The supernatant was decanted and pelleted nuclei were left in the inverted tubes on ice to drain briefly before being used in subsequent applications.

2.2.2 Nuclear protein isolation

The nuclei were resuspended in 1 ml of cold NE-1 buffer (20 mM HEPES pH 7.5, 10 mM KCl, 1mM $MgCl_2$, 0.1% Triton X-100, complete protease inhibitors [Roche] and 14.3 mM β-mercaptoethanol) and transferred to cold eppendorf tubes. After 5 min centrifugation at 600 g at 4 °C the supernatant was discarded and white nuclear pellets resuspended in 100 µl NE-1. 3 µl (250 units) of benzonase nuclease was added and the tubes were incubated for 10 min. at room temperature. Subsequently, volume was increased to 700 µl with NE-1 buffer and then 300 µl of NE-2 buffer (NE-1 supplemented with 500 mM NaCl) was added in a drop-wise fashion with mixing in order to uniformly reach 150 mM final concentration of NaCl. The nuclei were incubated 30 min on a rotating wheel at 4 °C. After centrifugation at 16 000 g for 20 min at 4 °C, the supernatant was transferred to new eppendorf tubes as a nuclear extract.

2.2.3 Protein electrophoresis

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed on home-made gels containing 0.1% SDS, 10-15% acylamide:bis-acrylamide (29:1; Bio-rad) in 375 mM Tris pH 8.8 (separating gel) or 4.3% acylamide:bisacrylamide in 125 mM Tris pH 6.8 (stacking gel), polymerised with ammonium persulphate and TEMED in Mini Protean-3 apparatus (Bio-rad). Protein samples were boiled in SDS PAGE sample buffer for 5 min and loaded on the gel alongside the PageRuler Prestained protein ladder (Thermo scientific). The electrophoresis was performed in SDS PAGE running buffer for 1 h at 180 V.

2.2.4 Western blotting

In order to transfer the proteins separated on the SDS-PAGE to nitrocellulose membrane, the wet transfer in Mini trans-blot cell system (Bio-rad) was used according to manufacturer's instructions. Briefly, the transfer cassette was assembled with sponge, 2 layers of Whatman paper, gel, membrane, 2 layers of Whatman paper and another sponge, all soaked in the protein transfer buffer. The transfer was

performed for 2 hours at 250 mA at 4 °C. The membrane was then blocked with 5% milk in TBS-T buffer (blocking solution) for at least 1 hour at room temperature, and incubated with primary antibodies diluted as in Table 2.1.2A in the blocking buffer at 4 °C overnight. Three washes for 10 min with TBS-T were followed by 2 hour incubation with secondary antibodies (Table 2.1.2B) diluted in blocking solution. Three washes as previously were followed by signal detection.

ECL detection

To visualise the proteins, membranes were incubated for 1 min in developing solution (0.1 M Tris-HCl pH 8.8, 0.01% H_2O_2, 1.5 mM luminol, 0.25 mM p-coumaric acid), drained and exposed to an X-ray film for 10 s – 30 min depending on the strength of the signal. The films were developed in SRX-101A film processor (Konica Minolta).

Ultra-sensitive ECL detection

In order to detect very low abundant proteins by Western blotting, the primary antibodies were used at higher than usually concentration (1:500 dilution), secondary antibodies were kept for longer (up to 4 hours) and to develop the signal a commercial SuperSignal WestFemto Chemiluminescent substrate (Thermo scientific) was used according to the manufacturer's protocol. Exposition onto the ultra-sensitive X-ray film (Hyperfilm ECL High Performance; Amersham) for 5 – 60 min was followed by film development as in classical ECL detection.

Quantitative infrared detection

To be able to quantitatively measure the intensity of Western blot bands the LiCOR infrared system was used. The membranes were processed with primary antibodies as in normal ECL detection, but were incubated with secondary antibodies coupled with fluorescent dyes (1:10 000 dilution) in the dark. The scanning of the signal and subsequent image analysis were performed on Oddysey scanner and Odyssey 3.0 software using the average background subtraction method.

2.2.5 Silver staining

Protein samples were resolved on the 10% polyacrylamide gel and stained with SilverQuest silver staining kit (Invitrogen) according to the manufacturer's instructions.

2.2.6 GFP immunoprecipitation

After nuclear protein extraction, 25 µl of the extract was mixed with 25 µl protein loading buffer, boiled for 5 min and stored at -20 °C as an input fraction. The remaining extract was mixed with 30 µl of GFP-Trap_A agarose beads (Chromotek), which have been previously equilibrated in NE-1 buffer supplemented with 150 mM NaCl (4 washes with 1 ml buffer, centrifugations performed for 1 min at 2 000 g). The extract was incubated with beads for 30 min on a rotating wheel in at 4 °C. The beads were pelleted by centrifugation (1 min, 2 000 g, 4 °C) and washed four times with 750 µl of buffer NE-1 supplemented with 150 mM NaCl, discarding the supernatants. The pelleted and washed beads were mixed with 60 µl protein loading buffer, boiled for 5 min and stored at -20 °C as the IP (immunoprecipitated) fraction for subsequent mass spectrometry or Western blot analyses.

2.2.7 Mass spectrometry

50 µl of IP samples were loaded onto NuPAGE pre-cast 4-12% Bis-Tris protein gels (Invitrogen). The electophoresis was run at 200V until the dye front migrated 1 cm into the gel. The gel was then washed 3 × 5 min with H_2O and stained for 1 h with the Imperial Protein coomassie stain (Thermo scientific), washed once with water and stored at 4 °C. Subsequent mass spectrometry analysis was performed by Flavia de Lima Alves in the laboratory of Juri Rappsilber. The bands of coomassie-stained proteins on the gel were excised and the proteins were digested with trypsin as described elsewhere (Shevchenko et al., 1996). In brief, proteins were reduced in 10 mM DTT for 30 min at 37 °C, alkylated in 55 mM iodoacetamide for 20 min at room temperature in the dark, and digested overnight at 37 °C with 13 ng/µl trypsin (Proteomics Grade, Sigma). The digestion media was then acidified to 0.1% of TFA

and spun onto StageTips as described in the literature (Rappsilber et al., 2003). Peptides were eluted in 20 µl of 80% acetonitrile in 0.1% TFA and were concentrated to 4 µl (Concentrator 5301, Eppendorf AG). The peptides sample was then diluted to 5 µl by 0.1% TFA for LC-MS/MS analysis. An LTQ-Orbitrap mass spectrometer (Thermofisher Scientific) was coupled on-line to an Agilent 1100 binary nanopump and an HTC PAL autosampler (CTC). The peptides were separated using an analytical column with a self-assembled particle frit (as described in: Ishihama et al., 2002) and C_{18} material (ReproSil-Pur C18-AQ 3 µm; Dr. Maisch GmbH) was packed into a spray emitter (100 µm ID, 8 µm opening, 80 mm length; New Objective) using an air-pressure pump (Proxeon Biosystems). Mobile phase A consisted of water, 5% acetonitrile, and 0.5% acetic acid; mobile phase B consisted of acetonitrile and 0.5% acetic acid. The gradient used was 98 min. The peptides were loaded onto the column at a flow rate of 0.7 µl/min and eluted at a flow rate of 0.3 µl/min according to the gradient. 0% to 5% buffer B in 5 min, 5% to 20% buffer B in 80 min and then to 80% buffer B in 13 min. FTMS spectra were recorded at 30,000 resolution and the six most intense peaks of the MS scan were selected in the ion trap for MS2 (normal scan, wideband activation, filling 7.5×10^5 ions for MS scan, 1.5×10^4 ions for MS2, maximum fill time 150 msec, dynamic exclusion for 150 sec). Identified peptides were searched against the mouse IPI protein library database (Kersey et al., 2004) and unique identifiers were assigned.

2.2.8 Mammalian cell culture

HeLa and NIH-3T3 cells were grown in DMEM medium supplemented with 10% foetal bovine serum (FBS; Hyclone), 100 U/ml penicillin and 100 µg/ml streptomycin. Cells were cultured in 75 cm^2 flasks (Corning) in 12 ml of complete medium at 37 °C, 5% CO_2 in standard cell culture incubator. When cells reached approximately 90% confluency they were washed with PBS, trypsinised with 1 ml TrypLE express trypsin and seeded at 1:10 dilution. For luciferase reporter gene assays and immunocytofluorescence cells were counted in haemocytometer and seeded at desired density.

2.2.9 Production of MeCP2 truncation mutants

In order to map the protein-MeCP2 interactions a series of FLAG-tagged truncation mutants of MeCP2 had to be produced. The primers containing NotI and BglII restriction enzyme sites were designed to amplify a desired region of human MeCP2_e2 and are available on request. The PCR reactions were performed in 25 µl final volume, using 50 ng pFS-MeCP2 plasmid as a template (a gift from M. Lyst), 2.5 µl 10× Pfu buffer (Promega), 200 µM dNTPs, 250 nM forward and reverse primers and 0.6 units of Pfu DNA polymerase (Promega). The reaction conditions were as follows: DNA denaturation at 95 °C for 1 min, then 30 cycles of 95 °C for 30 s; 60 °C for 30 s and 72 °C for 2 min, followed by final incubation at 72 °C for 10 min. The PCR product was purified on PCR purification column (Qiagen) according to manufacturer's protocol and eluted with 30 µl H_2O. The sample was double-digested with NotI and BglII restriction enzymes (in 40 µl total volume with 4 µl 10x buffer 2 [NEB], 1 mg/ml BSA, 1 µl NotI-HF and 2.5 µl BglII) for 2 hours at 37 °C. 5 µg of 3xFLAG-CMV-10 plasmid (Sigma) was digested in the same way. Both reaction solutions were mixed with Orange G DNA loading dye, separated on 1% agarose DNA gel with 0.5 µg/ml ethidium bromide in TAE buffer and the bands corresponding to digested vector and PCR product were excised with a scalpel under the UV light. The DNA was extracted from the gel with Gel extraction kit (Qiagen) and eluted with 30 µl H_2O. The ΔNLS2 mutant devoid of 255-269 aa fragment was produced from pFS-MeCP2 plasmid, using QuikChange II XL site directed mutagenesis kit (Stratagene) according to the manufacturer's instructions.

2.2.10 DNA ligation and plasmid amplification

The concentration of purified DNA was measured on Nanodrop spectrophotometer and 100 ng of vector was ligated with insert DNA at 1:3 molar ratio, using T4 DNA ligase (NEB) according to manufacturer's protocol. DH5α *E. Coli* bacteria were transformed with 10 µl ligation mixture by a standard heat shock method (Froger and Hall, 2007). The DNA from resulting colonies was extracted by Mini-prep method (Qiagen) and verified by diagnostic restriction digestion and sequencing to confirm

the correct length of desired truncation and lack of mutations. Plasmids were then amplified by Maxi-prep method (Qiagen), aliquoted and stored at -20 °C. Some of the MeCP2 truncation mutants were designed and produced by Matt Lyst.

2.2.11 DNA sequencing

Sequencing reactions were performed in 10 µl total volume, containing 2 µl BigDye Terminator v3.1 (Applied Biosystems), 4 µl 2.5 x DNA sequencing buffer, 5 pmol of sequencing primer and ~200 ng DNA. The reaction conditions were as follows: DNA denaturation at 96°C for 3 min, then 24 cycles of 96 °C for 30 s; 50 °C for 20 s and 60 °C for 4 min. Sequencing reactions were cleaned up and run on a ABI 3730 capillary sequencer by the GenePool sequencing facility (University of Edinburgh). Obtained sequences were analysed using BioEdit (Hall, 1999) and Lasergene v.10 software (DNAStar).

2.2.12 HeLa cell transfections for FLAG immunoprecipitation

For FLAG-immunoprecipitation experiments, ~5 × 10^6 HeLa cells were seeded on 15 cm dishes (Corning) in 22 ml of complete medium. 24 h after plating the cells were transfected with 20 µg of a plasmid harbouring the produced FLAG-tagged truncation mutant of MeCP2 or eGFP-C1 plasmid as a control, using JetPEI (Polyplus transfection) according to the manufacturer's protocol. The cells were harvested 24 h after transfection by trypsinisation, washed in PBS and stored at -80 °C.

2.2.13 FLAG immunoprecipitation

Harvested HeLa cells were defrosted on ice and homogenized in 1.5 ml cold NE-1 buffer in a 3 ml glass dounce. The nuclei were pelleted by centrifugation (5 min, 800 g, 4 °C) and the nuclear protein extraction was performed as in method 2.2.2. The FLAG immunoprecipitation was performed as in method 2.2.6, with the exception that 5 µl of washed anti-FLAG affinity agarose was used instead of GFP-Trap_A. After four washes of the beads the bound proteins were eluted with

20 µl of 1 mg/ml 3 × FLAG peptide (Sigma, F4799) for 3 hours on a rotating wheel in the cold room. Subsequently the beads were pelleted by 1 min centrifugation at 1000 g and the supernatants were mixed with 20 µl protein loading buffer, boiled for 5 min and stored at -20 °C as the IP (immunoprecipitated) fraction for subsequent Western blot analysis.

2.2.14 Cloning of CamKIIα gene

The primers were designed to amplify mouse CamKIIα gene (NCBI Reference Sequence: NM_177407.4) from mouse brain cDNA. Additionally, forward primer contained SalI restriction site and the reverse primer BamHI site and two linker nucleotides so the resulting product could be cloned in frame with GFP fusion protein in pEGFP-N1 plasmid (Clontech).

Table 2.2.14 Primers for cloning CamKIIα gene.

Primer name & direction	Restriction enzyme	Primer sequence 5'-3'
CamKII_FWD	SalI	ATTGTCGACATGGCTACCATCACCTGCACC
CamKII_REV	BamHI	TATGGATCCCCTCCATGCGGCAGGACG

Mouse brain cDNA was a gift from Pete Skene. The PCR reaction was performed and purified as in the method 2.2.9. Cloning to eGFP-N1 plasmid (Clonetech) was performed as in the method 2.2.10. For FLAG immunoprecipitations in HeLa cells 10 µg of the plasmid coding CamKIIα-GFP was co-transfected with 10 µg MeCP2 truncations and the IPs were performed as above.

2.2.15 Live cell imaging

Microscopic analysis and photography of live cells was performed on Nikon Eclipse TS100 inverted microscope in bright field and using mercury lamp and FITC filter for fluorescent images. The transfected ES cell-derived neurons were observed live under Zeiss Axiovert 25 inverted microscope fitted with mercury lamp and Filter sets 20 and 38.

2.2.16 Immunocytochemistry of MeCP2 expressed in NIH-3T3 cells

2×10^4 NIH-3T3 cells were seeded on each of the chambers of 8-chamber glass slide (BD Falcon) and transfected with 0.5 µg of GFP-MeCP2 or GFP-MeCP2 [255X] plasmids (gifts from L. Schmiedeberg) using JetPEI (Polyplus transfection) according to manufacturer's protocol. 24 h post-transfection the cells were washed 2 × 5 min with PBS, fixed 15 min with 4% paraformaldehyde (pH 7.4) in PBS, washed 3 × 5 min with PBS, permeabilised for 10 min with 0.1% Triton X-100 in PBS and washed as previously. Cells were then blocked with blocking solution (1.5% serum of the species where secondary antibodies were produced in PBS) for 30 min and incubated with primary antibodies (diluted in the blocking solution as listed in the Table 2.1.2A) for 2 h at room temperature. The 3 × 5 min washes with PBS were followed by incubation with secondary antibodies (Table 2.1.2B) for 1 h at room temperature in dark. Subsequently cells were washed 4 × 5 min with PBS and 4',6-diamidino-2-phenylindole (DAPI) was added to second wash (at 1 µg/ml final concentration). The coverslip was mounted with Prolong Gold (Invitrogen), corners were painted with nail varnish and slides were incubated at 4 °C overnight and stored at -20 °C.

2.2.17 Immunohistochemistry in mouse brain

The *Mecp2-GFP*$^{+/-}$ mouse was perfused with 4% paraformaldehyde (pH 7.4) in PBS (performed by H. Cheval and J. Selfridge) and the brain was removed, soaked in 4% paraformaldehyde overnight at 4 °C and in 30% sucrose for next 24 hours at 4 °C. Subsequently the brain was washed once with PBS, drained on paper, snap-frozen in isopentene on dry ice, moved to new container and stored at -80 °C. The brain slices were cut on CM1900 cryostat (Leica) at 14 µm thickness, placed on microscopic slides and frozen at -80 °C. The fluorescent immunostaining was performed as in method 2.2.16 from permeabilisation step, using coplin jars.

2.2.18 Picture analysis

The confocal microscopy was performed on Leica TCS SP5 microscope. The pictures were analyzed and processed in ImageJ v. 1.46r with LOCI tools plug-in (Schneider et al., 2012).

2.2.19 Biochemical fractionation of cytoplasmic and nuclear proteins

A single mouse brain was homogenised in 3 ml NE-1 buffer in 5 ml glass dounce by hand. After 15 min incubation on ice, 25 µl of the homogenate was boiled with equal amount of protein loading buffer and stored at -20 °C as a total brain protein extract. The rest of the homogenate was centrifuged for 10 min at 850 g at 4 °C. Supernatant was saved as cytoplasmic fraction and the pellets were washed twice with 1 ml NE-1 buffer (centrifuging as previously). The pellets was resuspended in 160 µl NE-1 buffer, 1 µl of benzonase nuclease was added and incubated at room temperature for 10 min. Subsequently, 240 µl NE-2 buffer was added to a 300 mM final NaCl concentration, vortexed and incubated 30 min on a rotating wheel at 4°C. After centrifugation at 16 000 g for 20 min at 4°C the supernatant was transferred to new tubes as a nuclear fraction.

2.2.20 Reverse-transcription quantitative PCR

The primers for gene expression analysis were designed in neighbouring exons of each gene, so the specific product of ~200 kb should be produced only from the cDNA, and possible contamination with genomic DNA would not give a product because of the presence of introns. Mouse brain cDNA used as a template in RT-qPCR was a gift from Pete Skene. Each reaction contained 10 µl 2x SensiMix SYBR & Fluorescein MasterMix (Bioline), 0.25 µM FWD primer, 0.25 µM REV primer, 5 µl of cDNA template and H_2O to 20 µl. The reactions were performed according to the following programme: initial denaturation at 95 °C for 10 min, 45 cycles of 95 °C for 10 s; 60 °C for 10 s 72 °C for 15 s (with an acquisition taken at the end of this step) in LightCycler 480 (Roche) and analysed in LightCyler480 Software 1.5.0.

Table 2.2.20 Primers for qPCR analysis of gene expression.

Primer	Sequence 5'-3'
Kpna3 FWD	AGTTGAGTGCTGTGCAGGCAGC
Kpna3 REV	TGTTTCCCAAAGCCCACACCGC
Kpna4 FWD	CCACAAAGACCCACCACCACC
Kpna4 REV	TGGTGGCTGAGTAGCGGAACC
Kpna6 FWD	AGCCGCGATGGAGACCATGG
Kpna6 REV	CCGGAATTTCTGCGTGGTTGC
Gapdh FWD	TACCCCCAATGTGTCCGTCG
Gapdh REV	CCTGCTTCACCACCTTCTTG

2.2.21 GAL4 luciferase reporter gene assays

4×10^4 of HeLa or NIH-3T3 cells per well were seeded on 24-well plates. In the initial experiments cells were co-transfected with JetPEI reagent with following amounts of plasmids: 1 µg GAL4 DBD-MeCP2 (C-terminal half, wild-type or R306C), 1 µg TK-Firefly (containing 5 GAL4 UAS sites) and 100 ng pRL-TK (control *Renilla* luciferase). We found that the use of limiting amounts of MeCP2 was critical to reveal the failure of repression by RTT mutants. Therefore, the optimised amounts of plasmids used in all subsequent assays were:
10 ng GAL4 DBD-MeCP2, 1 µg pEGFP-C1, 1 µg TK-Firefly, 100 ng pRL-TK.
Where indicated, 50 ng/ml of TSA (Sigma) was added to the medium at the moment of transfection. Each combination of plasmids was repeated in 2 - 4 wells. 48 hours post-transfection the cells were harvested and reporter gene expression was quantified using the Dual-luciferase reporter assay system (Promega) and TD-20/20 Luminometer. Transfection efficiencies were normalised by dividing the *Firefly* luciferase signal by *Renilla* luciferase levels. Fold repression of the *Firefly* luciferase reporter was calculated relative to a sample without MeCP2 or sample transfected with N-terminal half of MeCP2. The combined fold-repression was calculated by averaging fold repression of 2-6 biological replicates and the statistical significance of differences of repression by wild-type MeCP2 or mutated MeCP2 was assessed by unpaired two-sample Student's t-test assuming unequal variances.

2.2.22 ES cell culture

Mouse embryonic stem cells (section 2.1.4) were cultured in undifferentiated state in Glasgow MEM medium supplemented with 15% FBS, 1% sodium pyruvate, 1% non-essential amino acids, 0.1% β-mercaptoethanol, 100 U/ml penicillin, 100 µg/ml streptomycin and leukaemia inhibitory factor (LIF; gift from J. Guy) on gelatinised tissue culture plasticware. Cells were split by trypsinisation at 1:8 dilution every two days.

2.2.23 Neuronal differentiation system

The neuronal differentiation was performed as follows: 4×10^6 ES cells were plated on 10 cm Petri dish and cultured in 15 ml of ES medium with 10% FBS and without LIF. The media was changed every 2 days and 5 µM retinoic acid was added after 4 days. After 8 days the resulting embryoid bodies were treated with trypsin and cells were then resuspended in DMEM/F-12 medium with N2 supplement (Invitrogen) and passed through a 40 µm cell strainer (Falcon). 3×10^6 cells were plated on 6 cm dishes coated with poly-DL-ornithine hydrobromide (Sigma) and laminin (Roche) as described in (Bibel et al., 2004). After 24 hours the medium was replaced with a 50:50 mixture of N2 medium and Neurobasal medium with B27 supplement (Invitrogen). Every 3 days half of the medium was removed and replaced with Neurobasal/B27 medium. 6 days after plating cells were transfected for methylated luciferase assay or harvested for protein extraction. For immunocytochemistry the round 22 mm cover slips were put in the culture dishes. 6 days after plating the neuronal progenitors the cover slips were transferred to a 6-well plate and immunofluorescent staining was performed as in the method 2.2.16.

2.2.24 Transfection of ES cell-derived neurons

Test transfection of ES cell-derived neurons was performed in 24-well format to find the best gene delivery method. 1 µg of mCherry plasmid was transfected to each well containing 3×10^5 neurons. Five transfection reagents were tested using two volumes of reagent per well as follows: JetPEI (Polyplus transfection): 2 µl and 4 µl; Lipofectamine Plus (Invitrogen): 2 µl and 4 µl, both with 1 µl Plus reagent; X-treme

gene HP (Roche): 1 µl and 3 µl, X-treme gene 9 (Roche): 3 µl and 6 µl; 1 M $Ca_3(PO4)_2$: pH 7.0 and pH 7.1 (homemade), 6.2 µl; all according to the manufacturers' protocols. The best transfection method was chosen by live cell imaging, assessing the number of cells fluorescing red in each well.

2.2.25 Methylation of DNA

Methylation of 50 µg pRL-SV40 plasmid (Promega) was performed in 500 µl total reaction volume, containing 20 units of *M. SssI* CpG methyltransferase (NEB), 50 µl buffer 2 (NEB) and 2.5 µl S-adenosylmethionine (NEB). After 3 hours incubation at 37 °C, 2.5 µl of fresh S-adenosylmethionine was added and reactions were incubated for further 3 hours. As an unmethylated control the reactions were performed in the same way but without addition of the enzyme. The DNA was purified on PCR purification columns (Qiagen) and the efficiency of methylation was assessed by digestion of 200 ng DNA with methyl-sensitive restriction enzymes HpaII and HinP1I or methyl-insensitive enzyme MspI as a control for restriction reaction. The digestions were performed in 20 µl final volume with 1 µl of each enzyme and 2 µl of appropriate 10 × reaction buffer (NEB) for 2 hours at 37 °C. DNA was mixed with Orange G DNA loading buffer, separated on a 2% agarose gel.

2.2.26 Methylated reporter assay in neurons

The ES cell-derived neurons on 6 cm dishes were co-transfected with 1 µg of methylated or unmethylated pRL-SV40 plasmid, 1 µg pGL3 control plasmid (Promega) and 3 µg mCherry plasmid, using 15 µl X-treme gene 9 transfection reagent (Roche), according to the manufacturer's protocol. 48 hours after transfection the cells were harvested by scraping in 300 µl passive lysis buffer (Promega) and the reporter genes activities were quantified by three separate measurements using the Dual-luciferase reporter assay system (Promega) and TD-20/20 Luminometer. Transfection efficiencies were normalised by dividing the *Renilla* luciferase signal by control *Firefly* luciferase levels. Fold repression in each genotype of neurons was calculated by dividing the normalised reporter gene activity in cells transfected with non-methylated plasmid by normalised activity of cells with

methylated plasmid. Average fold repression was calculated from five independent biological replicates of this experiment and the statistical significance of differences between wild-type and MeCP2-mutated cells was measured by unpaired two-sample Student's t-test assuming equal variances.

Chapter 3: Identification of proteins binding to MeCP2

3.1 Introduction

One of the strategies applied in order to decipher the function of a protein is to identify the proteins it interacts with. This has been done for MeCP2 using various approaches, techniques and conditions, however the reports do not agree with each other. The initial attempt to identify binding partners of MeCP2 was based on the finding that it is able to repress transcription. Investigation of interactions with plausible candidate co-repressors revealed that MeCP2 binds to Sin3A complex (Jones et al., 1998; Nan et al., 1998). The complex contains histone deacetylases HDAC1 and HDAC2, which can explain why MeCP2-mediated transcriptional repression was alleviated by treatment with TSA. The fact that the repression was only partially relieved by TSA suggested that MeCP2 is able to regulate transcription by a different mechanism (Yu et al., 2000), possibly involving histone 3 lysine 9 methyltransferase (Fuks et al., 2003). A candidate protein driven search for other repressors binding to MeCP2 identified ample putative interactions: Ski, NCoR (Kokura et al., 2001); SMRT (Stancheva et al., 2003); CoREST, SUV39H1 (Lunyak et al., 2002); Brahma (Smarca2) (Harikrishnan et al., 2005, 2010); and cohesin subunits SMC1 and SMC3 (Kernohan et al., 2010). Also two transcription factors, TFIIB (Kaludov and Wolffe, 2000) and Sox2 (Szulwach et al., 2010) have been shown to bind to MeCP2 and possibly influence expression of other genes. Candidate approach revealed additional interactions of MeCP2 with proteins not directly involved in regulation of transcription: cyclin-dependent kinase-like 5 (CDKL5; Mari et al., 2005); DNA methyltransferase 1 (DNMT1; Kimura and Shiota, 2003); heterochromatin protein 1 (HP1; Agarwal et al., 2007); lamin B receptor (Guarda et al., 2009); and latency-associated nuclear antigen (LANA; Krithivas et al., 2004). However, the significance of these interactions *in vivo* has not been well established.

A different method of identifying proteins that possibly interact with MeCP2 was applied by Forlani and co-workers. In order to bind each other, proteins must be

expressed at the same time and space. A bioinformatic search of published microarray data for proteins fulfilling these conditions with MeCP2, revealed potential candidates and the interaction with one of them, a transcriptional regulator yin yang 1 (YY1) was confirmed (Forlani et al., 2010).

Unbiased, non-candidate driven searches for proteins interacting with MeCP2 have been performed only a few times. Young and co-workers over-expressed FLAG- and HA-tagged MeCP2 in HeLa cells and murine neuroblastoma cell line (Neuro2A) and subjected the co-immunoprecipitated material to mass spectrometry analysis (Young et al., 2005). Proteins were considered as specific binding partners of MeCP2 only when present in purifications with both tags from both cell lines. The authors did not publish the results of their search, but focused only on one identified protein, Y-box binding protein 1 (YB-1). Nevertheless, restricting the proteins interacting with mainly neuronal MeCP2 to the HeLa cell line raises a question about the physiological relevance of authors' findings, especially when considered from the Rett syndrome point of view.

A more physiological approach was taken by the same laboratory when MeCP2 binding partners were co-immunoprecipitated with anti-MeCP2 antibody from wild-type mouse brains and compared to *Mecp2*-null control mice (Chahrour et al., 2008). Interestingly, none of the proteins reported before to bind to MeCP2 were identified in this study. Notably low enrichment of MeCP2 by the antibody used could explain the small number of detected protein partners. Surprisingly, the first identified protein was a transcription factor CREB1, agreeing with authors' hypothesis that MeCP2 is an activator of transcription.

A different attempt to identify proteins interacting with MeCP2 used the yeast two-hybrid screen method. Because full-length MeCP2 caused transcriptional repression in yeast, the N-terminal half of MeCP2 was used as a bait for human foetal brain library (Nan et al., 2007). The screen detected MeCP2 binding to α thalassemia/mental retardation syndrome X-linked protein (ATRX), a member of the chromatin remodelling protein family. The binding was confirmed by other methods (Nan et al., 2007) and by other laboratories (Kernohan et al., 2010). It was also reported to be disrupted by a mutation in MeCP2 that causes X-linked mental retardation in males, A140V. Additionally, the ATRX localisation to

heterochromatic foci was altered in the *Mecp2* null cells (Nan et al., 2007). However, the Rett syndrome-causing mutations, even in the ATRX-interacting region did not affect the binding, putting in doubt the role of this interaction in RTT.

A yeast-two hybrid screen with the TRD domain of MeCP2 identified a homeodomain-interacting protein kinase 2 (HIPK2) that was shown to be able to phosphorylate MeCP2 (Bracaglia et al., 2009). However, the use of mouse embryo day 11 cDNA library in this study could mask other important interactions of MeCP2, given that it is expressed mainly in mature neurons.

MeCP2 has also been detected in screens for binding partners of other proteins. For example, it appeared on a list of interactors of PU.1 (Suzuki et al., 2003) and formin-binding protein 11 (FBP11), a splicing factor containing the WW domain (Bedford et al., 1997). This allowed discovery of a WW-domain binding region on MeCP2 at its C-terminus, responsible for this interaction (Buschdorf and Strätling, 2004). Another splicing factor identified to bind to WW-binding region of MeCP2 was huntingtin-interacting protein C (HYPC). Both interactions were abolished by C-terminal truncation of MeCP2, suggesting a potential role in the Rett syndrome pathology (Buschdorf and Strätling, 2004).

Despite two decades of research, using different approaches and techniques, there is little consensus about what are the true interacting partners of MeCP2 responsible for its function and its role in Rett syndrome pathology. Based on thorough biochemical analysis it has been reported that MeCP2 is in fact a monomer and does not stably associate with other proteins (Klose and Bird, 2004). However, that study used relatively high salt concentration to extract proteins from rat brain nuclei (400 mM NaCl) and could affect the weak interactions with partner proteins.

To date none of the published reports on interactions of MeCP2 described the effect of Rett syndrome mutations found frequently in the cluster at the end of TRD of MeCP2 (302-306). It is possible that these mutations disrupt general folding of the protein, affect its stability or prevent post-translational modifications that regulate its function. However, another possibility is that there are still some unidentified proteins interacting with MeCP2 and 302-306 mutations could disrupt their binding and in consequence cause RTT. Therefore, we decided to re-address the issue of

MeCP2 interacting partners, taking advantage of new molecular tools developed recently and perform an unbiased search for all of the proteins binding to the endogenous MeCP2 in mouse brains.

3.2 Identification of MeCP2-interacting proteins by mass spectrometry

In order to identify proteins interacting with MeCP2 we used a knock-in reporter mouse model recently obtained in the laboratory, expressing MeCP2 tagged with GFP. The animals were generated by J. Guy and J. Selfridge by inserting a GFP tag at C-terminal end of the protein by homologous recombination of *Mecp2* locus. As a result, the expression of MeCP2-GFP fusion protein is driven from the endogenous promoter of MeCP2 and levels of the protein should therefore be equal with the levels in wild-type mice. This notion was tested by extracting the nuclear proteins from brains of both animals and assessing the MeCP2 levels by Western blot. Wild-type MeCP2 and MeCP2-GFP were detected with comparable intensities demonstrating that the fused GFP tag does not affect the abundance of MeCP2 in the brain (Fig. 3.2A). There is, however, a risk that despite the equal levels of expression, the GFP tag can affect the function of a protein by altering its folding or disrupting its localisation due to increased size. The distribution of MeCP2-GFP has been assessed extensively by immunohistochemistry and confocal microscopy and was indistinguishable from the localisation of untagged MeCP2 (H. Cheval and J. Guy, personal communication; also see later chapters). The best indication that MeCP2-GFP is fully functional is that the transgenic mice are fertile, live till adulthood and their phenotypes appear to be indistinguishable from wild-type animals (J. Selfridge, unpublished). Moreover, an almost identical mouse model created by another group showed no detectable abnormalities and the expression profile and cellular localisation of MeCP2-GFP were alike the endogenous, untagged protein (Schmid et al., 2008). All of these observations prove that MeCP2-GFP mouse is a useful and valuable tool to study the function and characteristics of MeCP2 *in vivo*.

In order to identify binding partners of MeCP2, the nuclear proteins were extracted from brains of wild-type and MeCP2-GFP mice under the physiological salt concentration (150 mM NaCl). Such an ionic strength is not able to liberate MeCP2 bound to chromatin (Meehan et al., 1992), therefore the extracts were treated with benzonase nuclease. Both extracts contained comparable amounts of proteins and their composition was similar as assessed by a silver stained SDS-PAGE gel (Fig. 3.2B, input lanes). Subsequently, the proteins were subjected to immunoprecipitation (IP) with an anti-GFP peptide derived from llama (Rothbauer et al., 2008). The advantage of using this short binding peptide to pull-down the GFP-tagged protein is its small size (~13 kDa), which does not influence the mass spectrometric identification of proteins. Therefore, the contaminations from immunoglobulin heavy and light chains found in a typical antibody immunoprecipitation are avoided. Successful pull-down of proteins bound to MeCP2-GFP and almost no proteins precipitated from wild-type controls were verified on a silver-stained gel (Fig. 3.2B, IP lanes).

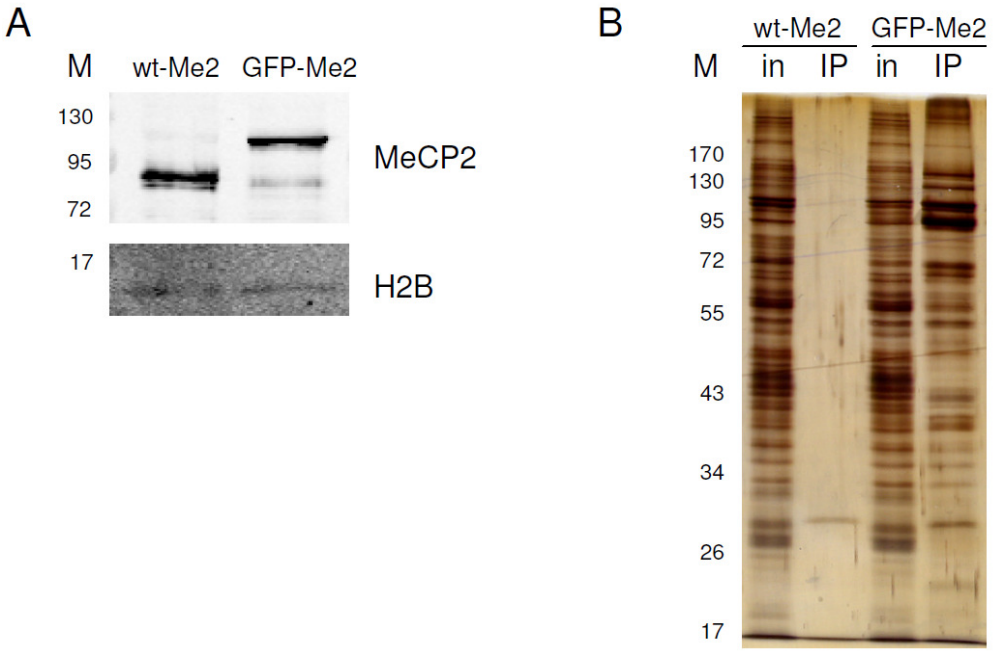


Figure 3.2. Immunoprecipitation of brain proteins from MeCP2-GFP mice. (A) Western blot showing comparable levels of MeCP2 in brain nuclei from wild-type (wt-Me2) and MeCP2-GFP knock-in mice (GFP-Me2). Histone H2B was used as a gel loading control. M – protein mass marker [kDa]. (B) Brain nuclear proteins were immunoprecipitated from wild-type mice or mice expressing MeCP2-GFP fusion protein with the anti-GFP peptide. in – 0.25% input, IP – 10% of immunoprecipitated material. The SDS-PAGE gel was silver-stained to visualise the proteins.

The immunoprecipitated proteins were subjected to identification by mass spectrometry (performed by F. de Lima Alves and J. Rappsilber). In total, 261 proteins were identified in both immunoprecipitations (all proteins are listed in Appendix Table 1). Amongst them 163 were unique for MeCP2-GFP IP, 63 were present in both purifications and 35 were found only in the wild-type brain IP. The high number of proteins identified in control IP is surprising, but mass spectrometry is such a sensitive method that even a small technical variation in IP efficiency or wash stringency could lead to a dramatic change in protein composition detected in each sample. Therefore, in order to produce a comprehensive list of plausible proteins interacting specifically with MeCP2 the results were compared to a biological replicate of the same experiment (performed by M. Lyst). Only the proteins represented by at least 2 peptides in both purifications from MeCP2-GFP brains, but not present in any of the wild-type MeCP2 control IPs, were considered as candidate specific interactors of MeCP2. Such a filtered list contains 40 proteins (Appendix Table 2). The top 25 of putative proteins associated with MeCP2 arranged by a number of identifying peptides are listed in the Table 3.2.

Surprisingly, the obtained list contains only two proteins that have been reported previously to interact with MeCP2, namely NCoR1 (Kokura et al., 2001) and SMRT (Stancheva et al., 2003). The other proteins suspected to be interacting partners of MeCP2 (see section 3.1) were not present in our purification. Sin3A, considered as one of the main co-repressors associated with MeCP2 and thought to be responsible for its function in suppressing transcription (Jones et al., 1998; Nan et al., 1998) was represented only by one peptide in my purification and therefore was excluded from the list (see Appendix). Mass spectrometry analysis of the biological replicate of this experiment detected 3 peptides of Sin3A, suggesting that this interaction is not the most prominent, when compared to the proteins from the top of the list. Other components of the Sin3A co-repressor complex were also detected by only few peptides (HDAC1) or were present in one of the wild-type control purifications (HDAC2). Although our mass spectrometry analysis was not quantitative and many factors could influence the number of recovered peptides (such as the efficiency of in-gel digestion, extraction, and peptide separation), the nature of the association of MeCP2 with Sin3A was reported before as weak and

unstable (Klose and Bird, 2004), what was additionally confirmed by immunoprecipitation from MeCP2-GFP brains followed by Western blot analysis (M. Lyst, personal communication).

Table 3.2. Top 25 proteins identified in GFP immunoprecipitations from MeCP2-GFP brains. Proteins are arranged according to the number of identifying peptides in my experiment (Exp. #1). Column labelled 'Exp. #2' contains number of peptides assigned to each protein in the biological replicate of the experiment. Proteins highlighted in green belong to one co-repressor complex and were investigated further in the laboratory. Interactions of MeCP2 with proteins highlighted in yellow were confirmed and followed-up in this study. All of the proteins identified by mass spectrometry and a complete filtered list can be found in the Appendix Table 1 and Appendix Table 2.

Peptides		Protein coding gene	Protein name
Exp. #1	Exp. #2		
52	57	*Mecp2*	Methyl-CpG-binding protein 2, Isoform B
27	5	*Ncor2*	Nuclear receptor co-repressor 2 (SMRT)
20	9	*Ncor1*	Nuclear receptor co-repressor 1
17	13	*Kpna4*	Importin subunit α4 (KPNA4)
16	12	*Tbl1xr1*	Transducin β-like 1 X-related protein 1 (TBLR1)
12	15	*Eif4a3*	Eukaryotic initiation factor 4A-III
11	7	*Cdc5l*	Cell division cycle 5-related protein
11	19	*Prpf19*	Pre-mRNA-processing factor 19, Isoform 1
10	7	*Kpna3*	Importin subunit α3 (KPNA3)
8	3	*Camk2a*	Calcium/calmodulin-dependent protein kinase II α
8	2	*Ncoa5*	Nuclear receptor coactivator 5
8	2	*Raly*	RNA-binding protein Raly, Isoform 2
7	10	*Rbbp4*	Retinoblastoma binding protein Rbbp4
7	10	*Tra2b*	Transformer-2 protein homolog β, Isoform 1
7	7	*Sfrs7*	Splicing factor, arginine/serine-rich 7, Isoform 2
7	6	*Snrnp70*	U1 small nuclear ribonucleoprotein 70 kDa, Isoform 1
7	5	*Hdac3*	Histone deacetylase 3
7	5	*Tbl1x*	Transducin β-like 1 X-linked (TBL1)
6	17	*Smarcc2*	SWI/SNF complex subunit Smarcc2, Isoform 2 (BAF170)
5	2	*Snw1*	SNW domain-containing protein 1
5	2	*Wbp11*	WW domain-binding protein 11
4	9	*Csnk2a1*	Casein kinase II subunit α (CK2α)
4	9	*Wdr18*	WD repeat-containing protein 18
4	5	*Magoh-rs1*	Mago-nashi homolog
4	5	*Smu1*	WD40 repeat-containing protein SMU1, Isoform 1
4	4	*Fyttd1*	Forty-two-three domain-containing protein 1, Isoform 1

Some of the newly identified proteins drew our attention. Interestingly, five out of top sixteen putative MeCP2-interacting partners (SMRT, NCoR1, TBLR1, TBL1 and HDAC3; marked green in Table 3.2) are members of NCoR/SMRT co-repressor complex. These interactions may have important consequences for the MeCP2 function as a transcriptional repressor and were investigated further (see Chapter 5). Two other mouse proteins at the top of the list, importin subunit α3 (also known as karyopherin α3; KPNA3) and importin subunit α4 (karyopherin α4; KPNA4) belong to the importin α family and interactions with them may explain how MeCP2 is transported to the nucleus (see Chapter 4). The final two candidate proteins picked for subsequent analysis, calcium/calmodulin-dependent protein kinase II α (CaMKIIα) and casein kinase 2α (CK2α) are protein kinases and could possibly regulate the function of MeCP2 through phosphorylation.

3.3 Verification of mass spectrometry data by Western blotting

Mass spectrometry is a very sensitive technique and the results have to be interpreted cautiously to avoid false positives. Therefore, it is important to verify every interaction by means of other methods. To validate the selected subset of discovered binding partners of MeCP2, proteins were extracted from new MeCP2-GFP mouse brain nuclei and immunoprecipitated with anti-GFP antibody, as previously. As a control, the brain nuclear extracts from mouse expressing H2B-GFP fusion protein were used (Jiang et al., 2008). Resulting IP samples and their corresponding inputs were analysed by Western blot, probing with antibodies raised against each of the tested putative MeCP2-interacting protein.

3.3.1 Validation of interactions between MeCP2 and importins

In order to control the feasibility of the whole procedure the Western blot membranes were firstly incubated with anti-MeCP2 antibody. MeCP2 and its GFP-tagged version were detected in the inputs from brains of both genotypes confirming that they were expressed at comparable levels (Fig. 3.3.1A and B). The bands corresponding to MeCP2 were not detected in the immunoprecipitation lanes

of H2B-GFP mice, but as expected, they were present only in the IP lanes from the MeCP2-GFP samples (Fig. 3.3.1A and B). These results proved the successful immunoprecipitation of MeCP2-GFP with anti-GFP binding peptide and the validity of the method.

Probing the same Western blot membranes with antibodies raised against importins α detected KPNA3 and KPNA4 in the input lanes from both mice genotypes, but only in the MeCP2-GFP IP lanes, not in the H2B-GFP IP (Fig. 3.3.1C and D). The co-immunoprecipitation of KPNA3 and KPNA4 with MeCP2 suggests that the importins identified by mass spectrometry are indeed specific interacting partners of MeCP2. Both interactions seemed very strong as quantitative densitometric analysis of IP to input ratios revealed that approximately 40% of all KPNA3 was co-immunoprecipitated with MeCP2 (Fig. 3.3.1C). The binding of KPNA4 to MeCP2 was even stronger, reaching up to 93% of total protein (semi-quantitative Western blot; Fig. 3.3.1D).

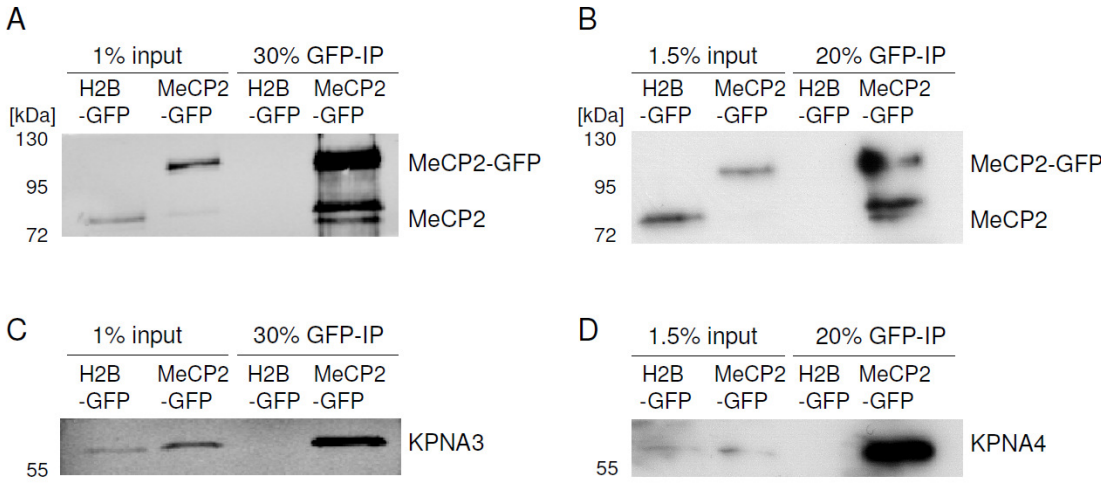


Figure 3.3.1. Verification of MeCP2 interaction with importins. Brain nuclear proteins were immunoprecipitated with GFP binding peptide (GFP-IP) from mice expressing H2B-GFP or MeCP2-GFP fusion proteins and analysed by Western blot. Percentage numbers indicate the fraction of each sample loaded on the gel. The figure shows representative pictures of two independent experiments. (A and B) Detection of MeCP2 and MeCP2-GFP by anti-MeCP2 antibody. (C) Probing the same Western blot membrane as A for KPNA3. (D) Probing the same Western blot membrane as B for KPNA4.

3.3.2 Validation of interactions between MeCP2 and kinases

Other proteins supposedly binding to MeCP2 and selected for validation were CK2α and CaMKIIα protein kinases. To verify the data obtained with mass spectrometry, immunoprecipitations were performed as previously from new brain nuclear extracts of MeCP2-GFP and H2B-GFP mice. Subsequent Western blot analysis revealed that both kinases were co-immunoprecipitated only from MeCP2-GFP brains, confirming the interactions between MeCP2 and CK2α and CaMKIIα (Fig. 3.3.2A and B, respectively). Quantification of the proteins detected in the IP fractions and in the inputs showed that MeCP2 binding to kinases was much weaker than in the case of importins. Because of the low signal and high background from the antibodies used, the ratios could be only approximately estimated and ~0.32% of all CK2α was bound by MeCP2 (Fig. 3.3.2A) and just ~0.22% of all CaMKIIα (Fig. 3.3.2B).

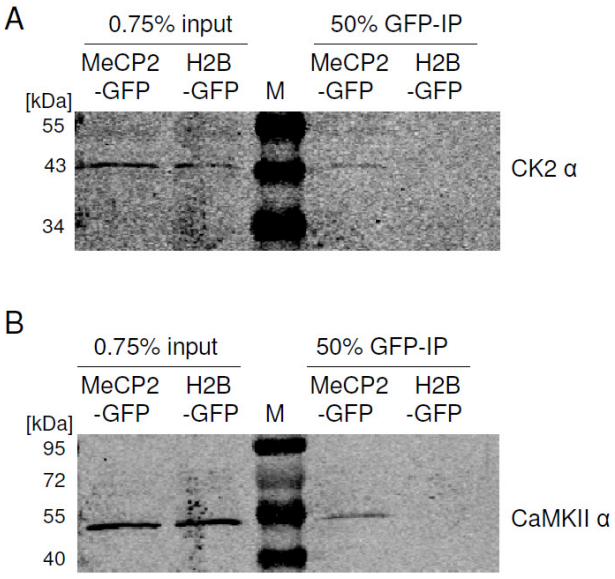


Figure 3.3.2. Verification of interactions of MeCP2 with kinases. Brain nuclear proteins were immunoprecipitated with GFP binding peptide from mice expressing H2B-GFP or MeCP2-GFP fusion proteins and analysed by Western blot probing for (A) casein kinase 2α (CK2α) or (B) calcium/calmodulin-dependent protein kinase II α (CaMKIIα). These two experiments were performed twice with help of honours student S. Young under my supervision.

3.3.3 Validation of other interactions of MeCP2

The interactions between MeCP2 and identified components of NCoR/SMRT co-repressor complex (NCoR1, SMRT, HDAC3 and TBLR1) were confirmed in a similar fashion in the laboratory (M. Lyst, personal communication). The interactions of MeCP2 with these proteins also appeared to be weaker than with importins.

Some other proteins identified by mass spectrometry (WBP11, MAP1B) were preliminarily tested for interaction with MeCP2, but could not be detected by Western blot analysis in the immunoprecipitates (data not shown). This may indicate either low sensitivity of the antibodies used or the false positive hits in the mass spectrometry analysis.

3.4 Mapping the confirmed interactions on the MeCP2 sequence

3.4.1 Generation of MeCP2 truncation mutants

The next step in analysis of interaction between two proteins is mapping the region of their binding. This can be achieved by expressing fragments of the tested proteins and assessing their binding. A series of FLAG-tagged truncation mutants of MeCP2, spanning its different regions and domains, was designed and produced. To allow the assessment of the potential influence of Rett syndrome mutations on the identified interactions, the MeCP2 constructs were designed with respect to the clusters of RTT mutations (Fig. 3.4.1-1). Each of the produced expression construct was verified by DNA sequencing.

The constructed MeCP2 fragments were then expressed in HeLa cells. Transfection efficiency was estimated by GFP signal in cells transfected with a plasmid expressing EGFP protein and was between 60% and 70% (Fig. 3.4.1-2; analysis of four independent transfections). Subsequent immunoprecipitations with anti-FLAG beads were followed by Western blot analysis to assess which fragments of MeCP2 bind the tested protein of interest. Each membrane was re-probed with anti-FLAG antibody to ensure the equal expression of used truncations of MeCP2.

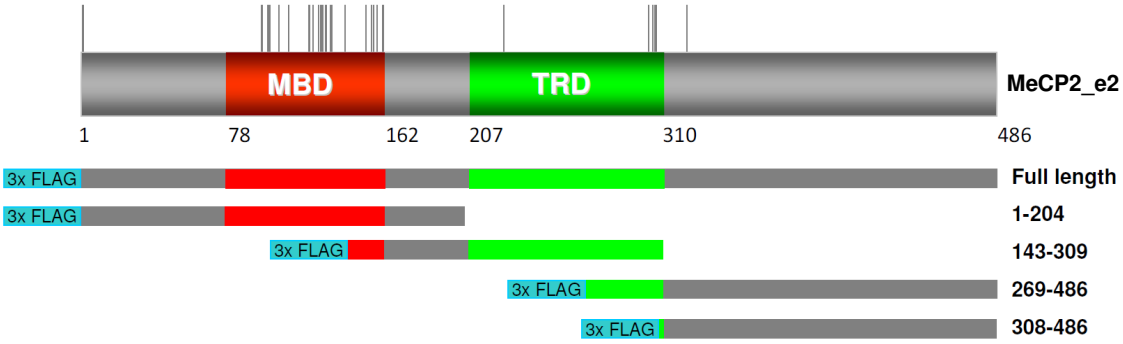


Figure 3.4.1-1. Design and production of MeCP2 truncation mutants. The truncation mutants of MeCP2 were amplified by PCR from MeCP2_e2 human cDNA and cloned in frame into a vector expressing 3×FLAG tag. A schematic drawing of just a few of the produced constructs used in this study. The numbers correspond to amino acid positions in MeCP2_e2. Mutations causing Rett syndrome are depicted as vertical lines above MeCP2. Some of the clones were designed and produced by M. Lyst.

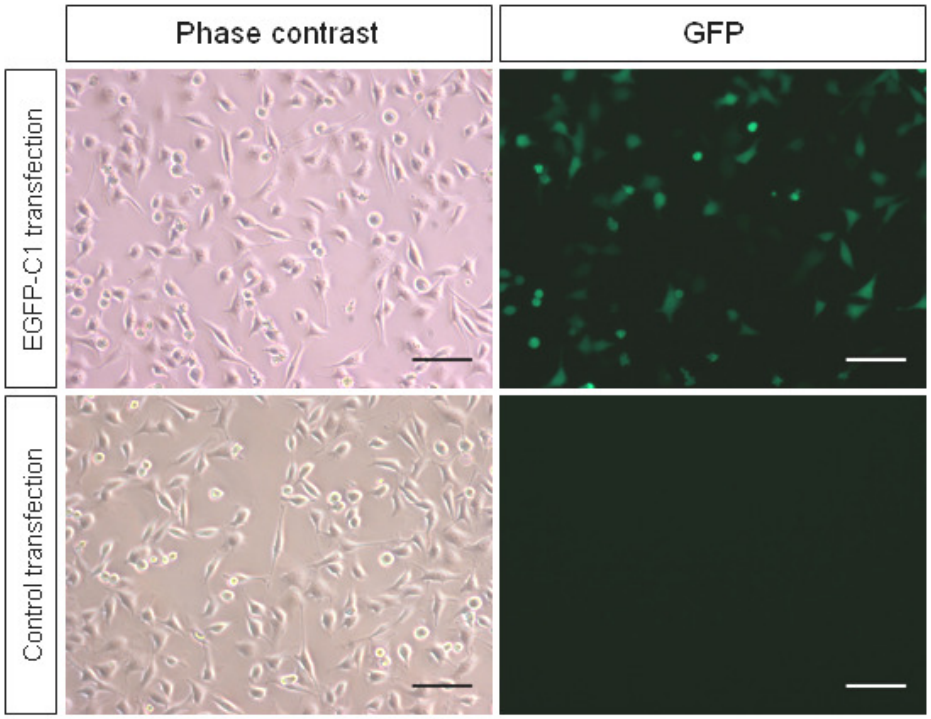


Figure 3.4.1-2. HeLa cells transfection efficiency for immunoprecipitations. HeLa cells were transfected with EGFP-C1 plasmid and photographed live under the fluorescent microscope. Green fluorescent cells were counted and their number divided by the total number of cells counted using the phase contrast miscroscopy. Representative pictures of four independent transfections. Scale bar: 50 µm.

3.4.2 Mapping the interactions of MeCP2 with importins α and subunits of NCoR/SMRT co-repressor complex

Using the methodology described above the interactions between MeCP2 and importins α (KPNA3 and KPNA4) were extensively mapped to a specific region. The strategy, design of new MeCP2 truncations, results and further detailed analysis of the role of these interactions are described in the Chapter 4.

Identification of the region of MeCP2 interacting with the subunits of NCoR/SMRT co-repressor complex was performed in a similar fashion by M. Lyst and the obtained results and their consequences follow in the Chapter 5.

3.4.3 Mapping the interaction of MeCP2 with casein kinase 2α

Casein kinase 2 is a promiscuous protein kinase expressed across a range of eukaryotic tissues and cells, including HeLa cell line (Edelman et al., 1987; Yu et al., 1991). Therefore, binding of the endogenous protein to over-expressed fragments of MeCP2 could be tested. Immunoprecipitation of FLAG-tagged full length MeCP2 and subsequent probing with anti-CK2α antibody confirmed that these proteins interact with each other in HeLa cells (Fig. 3.4.3, top). From two bands detected in all of the input fractions only the bottom one was immunoprecipitated. Our prediction is that the top band is a protein in the HeLa extract detected unspecifically by the antibody (it migrates at higher size than predicted size of CK2α) but is has not been tested. Surprisingly, analysis of interactions of truncated versions of MeCP2 with the kinase revealed that each of the tested MeCP2 fragments bound CK2α. The strongest signal was detected in co-immunoprecipitations with full-length MeCP2 and its non-overlapping fragments, 1-204 and 308-486 (Fig. 3.4.3, top). The other tested truncations of MeCP2 (143-309, 202-486 and 269-486) also weakly interacted with the kinase, suggesting that the whole length of MeCP2 can act as a CK2α binding surface. However, there was also a very weak CK2α band detected in a control IP lane, where immunoprecipitation was performed with the anti-FLAG beads in HeLa cells transfected with EGFP plasmid only (Fig. 3.4.3, top). This implies that at least some of the CK2α binding detected in all IP lanes may be due to the unspecific interactions with anti-FLAG beads.

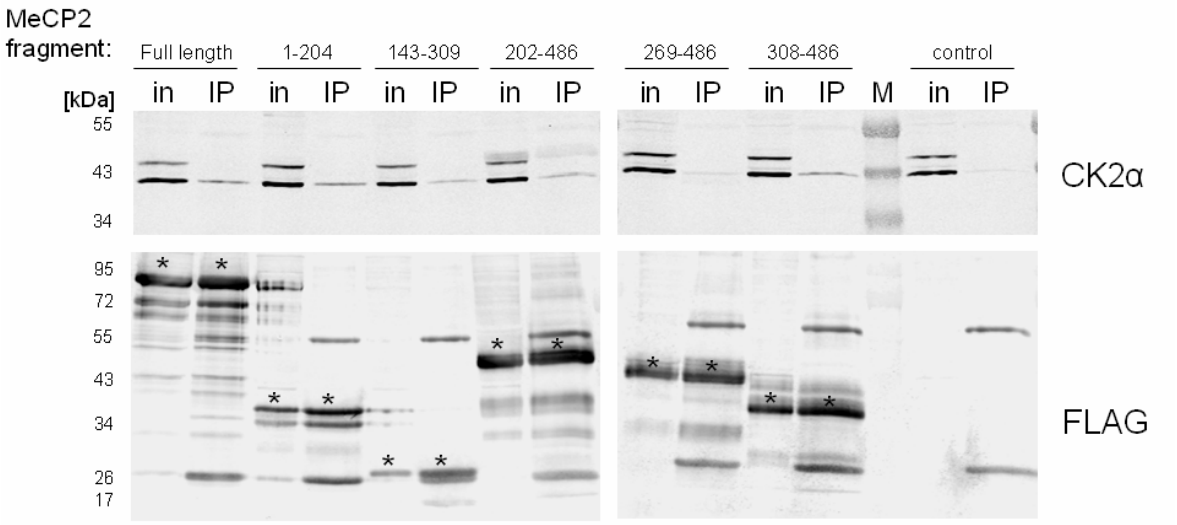


Figure 3.4.3. MeCP2 interacts with CK2α at the whole length of the protein. HeLa cells were transfected with FLAG-MeCP2 truncation mutants or GFP-expressing plasmid (control). Inputs (in) and FLAG-immunoprecipitations (IP) were resolved on SDS-PAGE, Western blotted and probed for CK2α. Re-probing of the same membrane with anti-FLAG antibody. This experiment was performed once by honours student S. Young under my supervision.

All of the tested MeCP2 truncations were expressed and immunoprecipitated at similar levels, as assessed by re-probing the Western blot membrane with anti-FLAG antibody (Fig. 3.4.3, bottom; asterisks).

3.4.4 Mapping the interaction of MeCP2 with CaMKIIα

Another kinase found in this study to interact with MeCP2 was calcium/calmodulin-dependent protein kinase II α. Western blot analysis revealed that it is not expressed in HeLa cells (data not shown). Therefore, to be able to perform the mapping analysis it had to be co-expressed with MeCP2 truncations. *CamkIIα* mouse gene was amplified from mouse brain cDNA, purified and cloned into a GFP-expressing plasmid to produce a fusion protein, which could be then easily tracked in cells. The N-terminus of CaMKIIα contains a conserved domain responsible for binding of kinase substrates, and we hypothesised that it could be a potential interaction surface for MeCP2. The C-terminus of CaMKIIα consists of a less conserved self-association domain (Hoelz et al., 2003; Chao et al., 2011). Therefore, to avoid the potential disruption of CaMKIIα binding surface, the GFP tag

was placed at the C-terminus of the protein, where the fused protein was less likely to affect the interaction with MeCP2.

The co-expression of CaMKIIα-GFP and full length FLAG-MeCP2 in HeLa cells, subsequent FLAG-immunoprecipitations and Western blot analysis revealed that these two recombinant proteins bind to each other (Fig. 3.4.4A). Interestingly, a band migrating higher than the predicted molecular weight of CaMKIIα-GFP was detected in the input fractions and was strongly enriched in the FLAG-MeCP2 IP (Fig. 3.4.4A). It is possible that this band is a post-translationally modified version of CaMKIIα and it is preferred by MeCP2, however the character of this band has not been determined. The interaction was also detected between CaMKIIα-GFP and MeCP2 1-204 fragment, and, to a lower extent, with the 143-309 fragment. The N-terminal truncations of MeCP2 (269-486 and 308-486) did not co-immunoprecipitate CaMKIIα. The FLAG-IP procedure in control cells, expressing CaMKIIα-GFP, but no FLAG-MeCP2, also failed to precipitate the kinase.

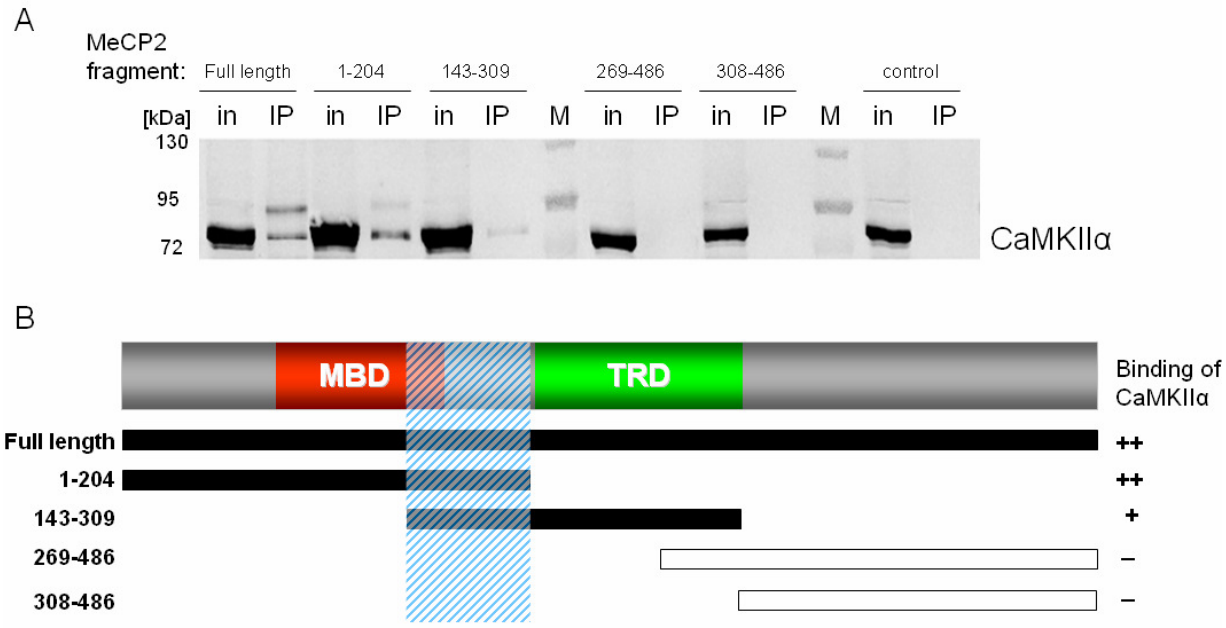


Figure 3.4.4. MeCP2 interacts with over-expressed CaMKIIα at 143-204. HeLa cells were co-transfected with plasmids expressing CaMKIIα-GFP and FLAG-MeCP2 truncation mutants or only CaMKIIα-GFP (control). (A) Inputs (in) and FLAG-immunoprecipitations (IP) were resolved on SDS-PAGE gel, Western blotted and probed for CaMKIIα. These experiments were performed twice with help of honours student S. Young under my supervision. (B) Mapping of the MeCP2-CaMKIIα interaction on the MeCP2 sequence with a minimal interacting region marked with blue dashed box.

Summarizing, it seems that CaMKIIα interacts with the N-terminal part of MeCP2, within a minimal region mapped between amino acids 143-204, and probably some additional binding sites between 1-143 (Fig. 3.4.4B). Interestingly, this region overlaps with the MBD, a known hotspot for Rett syndrome causing mutations. The effect of these mutations on the ability of MeCP2 to interact with CaMKIIα should be tested in the future to reveal if there is any correlation between this interaction and pathology of Rett syndrome.

3.5 Discussion

Immunoprecipitation of MeCP2 from mouse brains and identification of co-immunoprecipitated proteins by mass spectrometry performed by us in this study provided a new list of potential protein partners of MeCP2. Out of 40 identified proteins, only two of them have been reported previously. The reason for the discrepancies between ours and published results may lie in the technical approach taken by each study. Some of the previously performed purifications of MeCP2 used high salt concentrations (300-400 mM NaCl) what could be too stringent for most of the interactions (Lunyak et al., 2002; Klose and Bird, 2004; Agarwal et al., 2007; Gonzales et al., 2012). Other studies suffered from weak affinity of anti-MeCP2 antibody (Chahrour et al., 2008). Thanks to the recently developed new molecular tools we were able to perform efficient co-immunoprecipitations under the physiological salt conditions. We were aware that lower stringency of protein extraction and washes is burdened with a risk of retaining some weak unspecific interactions, which could be detected by sensitive mass spectrometry. However, analysis of two independent protein purifications allowed us to form a relatively short list of plausible MeCP2 interactors. Eight out of nine proteins chosen for verification by Western blotting confirmed the mass spectrometry results. Other evidence for the appropriateness of our approach is that some of the repeatedly confirmed interactions of MeCP2 present in our purifications are abolished at 300 mM NaCl (M. Lyst, personal communication).

The fact that some of the interactions reported before were not found under our conditions may indicate that they are not the major partners of MeCP2 and were detected only because of the candidate approach or as artefacts. An example of Sin3A is the best illustration of the former. As MeCP2 was reported to be able to repress transcription a quest for finding the repressor interacting with it was taken by many laboratories. Sin3A, an HDAC-containing co-repressor complex was one of the plausible candidates and indeed it was co-immunoprecipitated with MeCP2 (Jones et al., 1998; Nan et al., 1998). Although the biochemical analysis showed that this interaction is weak and not stable (Klose and Bird, 2004) it was considered to be the repressor responsible for the main function MeCP2. Also in our study the binding of Sin3A and the subunits of the co-repressor, HDAC1 and HDAC2 to MeCP2 was weak (based on number of identifying peptides as well as Western blot analysis performed by M. Lyst). However, we have identified most of the components of another co-repressor complex, NCoR/SMRT. The abundance of identified peptides of each of the protein of the complex in our both co-immunoprecipitations was much higher than for Sin3A complex. Moreover, the candidate approach studies also revealed the interactions between MeCP2 and NCoR1 and SMRT (Kokura et al., 2001; Stancheva et al., 2003). As the function of both Sin3A and NCoR/SMRT complexes in terms of transcriptional repression is similar, our data raises possibility that NCoR/SMRT can be the major co-repressor binding to MeCP2 and contributing to the transcriptional repression function of MeCP2 more than Sin3A. This notion has been addressed and is described in Chapter 5.

Interestingly, one of the core components of NCoR/SMRT co-repressor complex, G protein pathway suppressor 2 (GPS2), was not identified in any of our mass spectrometry screens. In addition, Western blot analysis failed to detect GPS2 in GFP immunoprecipitations from mouse brains expressing MeCP2-GFP (M. Lyst, personal communication). It has been shown that GPS2 interacts with both TBL1 and SMRT (Oberoi et al., 2011), therefore it is possible that MeCP2 is competing for its binding sites and these proteins are mutually exclusive in the complex. Another possibility is that due to its small size (36.6 kDa) GPS2 was missed in mass spectrometry identification and the antibodies used for Western blot were not sensitive enough to detect the indirect interaction. This issue remains unanswered at

the moment, but should be addressed in the future with additional antibodies or *in vitro* binding assays.

Another explanation for the discrepancies between the published interactions of MeCP2 and the data obtained in our purifications may be that some of the interactions may be activity-dependent. Since the phosphorylation of MeCP2 has been shown to be regulated by neuronal activity (Zhou et al., 2006; Tao et al., 2009), it is possible that some of the interactions with MeCP2 may be disrupted and others transiently present only upon neuronal activation. The preference of some of the candidate protein partners of MeCP2 to bind it when phosphorylated was recently reported (Gonzales et al., 2012). The unbiased investigation of the composition of proteins co-immunoprecipitating with MeCP2 upon seizure induction in mice would be an interesting set of experiments shedding more light onto this issue.

Amongst the proteins identified by us as interacting with MeCP2 there are also other candidates of potential interest. For example, retinoblastoma binding protein 4 (RBBP4 also known as RbAp48) is a histone binding protein that can bring histone substrates to the complexes involved in chromatin remodelling or histone deacetylation (Nicolas et al., 2000). It has been found to be associated with NuRD and Sin3 co-repressor complexes (Sun et al., 2007), as well as with nucleosome remodelling factor complex (NURF; Barak et al., 2003). Another protein implicated in chromatin remodelling and putatively interacting with MeCP2 is chromodomain helicase DNA binding protein 4 (CHD4). It was additionally suggested to be involved in histone deacetylation (Schmidt and Schreiber, 1999). A different protein putatively binding to MeCP2 and being a part of SWI/SNF chromatin remodelling complex is a member 2 of subfamily C of SWI/SNF related, matrix associated, actin dependent regulator of chromatin (SMARCC2, also known as BAF170). A protein belonging to the same complex (SMARCA2, also known as Brahma) has been identified to bind to MeCP2 before (Harikrishnan et al., 2005, 2010). Although it was not detected in our present purifications and the validity of its interaction with MeCP2 was questioned previously (Harikrishnan et al., 2006; Hu et al., 2006), the fact that yet another component of the complex seems to be co-immunoprecipitated with MeCP2 suggests that this elusive interaction, may in fact be real. The

importance of these findings remains to be determined. All of the proteins mentioned above can contribute to the role of MeCP2 as transcriptional repressor, and the character and significance of the interactions between them should be investigated in the future.

We have also confirmed and investigated further the interactions of MeCP2 with importin α family proteins, KPNA3 and KPNA4. The quantification of detected binding revealed that these interactions are very strong, with 40% of KPNA3 and more than 90% of KPNA4 molecules bound by MeCP2 in brains. These values seem to be unbelievable high and possibly do not reflect the strength of a real *in vivo* interaction, but may at least partially be due to the high affinity of importins to the exposed NLS of MeCP2 in the homogenised nuclei. However, even if the real interactions are weaker than presented here, they are still much stronger than of the other analysed interactions, suggesting their significance. The functionality, consequences and the relationship between MeCP2 and KPNA3 and KPNA4 are described in Chapter 4.

The interactions of MeCP2 with subunits α of two serine/threonine protein kinases identified by mass spectrometry were confirmed by us using another method and mapped on the MeCP2 sequence. Both CK2 and CaMKII are abundant in neurons and have been implicated in regulation of synaptic plasticity, neurite extension and memory formation (Lieberman and Mody, 1999; Yamauchi, 2005; Lucchesi et al., 2011).

According to the preliminary data CK2α binds to the whole length of MeCP2. It is possible that CK2 has multiple binding sites on MeCP2, similar to the reported interaction of MeCP2 with for example for Sin3A (Nan et al., 1998). Another explanation is that the detected interaction is indirect and mediated by other proteins binding to different regions of MeCP2. Given a broad range of protein substrates of CK2 (Pinna, 2002; Meggio and Pinna, 2003) it is a hypothesis, that should be formally tested. The role of CaMKII in activity-dependent phosphorylation of serine 421 of MeCP2 has been reported before (Zhou et al., 2006), but it was later suggested that it is rather mediated by another kinase, CaMKIV (Tao et al., 2009).

Notably, the interaction between MeCP2 and these proteins *in vivo* has never been demonstrated and it could be that they act indirectly in a cascade of kinases. Two kinases reported to interact with MeCP2 and mediate its phosphorylation, CDKL5 (Mari et al., 2005) and HIPK2 (Bracaglia et al., 2009) have not been detected in our purifications. However, our discovery brings CaMKII kinase back in the spotlight as a potential regulator of MeCP2 function. The identified region of interaction (143-204, with possible additional binding sites in the N-terminus of MeCP2) contains threonine 148 and serine 149, identified as phosphorylated residues of MeCP2 (Tao et al., 2009; Gonzales et al., 2012). It is also in proximity to other reported sites of MeCP2 phosphorylation: serines 80, 164 and 229. Although the most studied phosphorylated serine 421 lies further away from the putative CaMKII binding region, the very complex supramolecular organisation of the kinase, consisting of 14 subunits, may allow it to interact in a distance from the docking site (Hoelz et al., 2003). Interestingly, the putative CaMKII interacting region overlaps with the MBD domain, where numerous Rett syndrome causing mutations are found, including the most frequent, T158M. The effect of these mutations on the ability of MeCP2 to bind the kinase should be investigated and may give some evidence on the aetiology of RTT.

One of the alleged functions suggested for MeCP2 is transcriptional activation (Yasui et al., 2007; Chahrour et al., 2008; Ben-Shachar et al., 2009). The putative protein postulated to be involved in this process, a transcription factor CREB1, has not been identified in our co-immunoprecipitations. However, the mass spectrometry list of MeCP2 interactors contains other proteins that may potentially act as activators of transcription. SNW domain-containing protein 1 (SNW1) is able to activate vitamin D receptor and retinoic acid receptor (Baudino et al., 1998; Kang et al., 2010). Nuclear receptor co-activator 5 (NCoA5) also known as co-activator independent of AF-2 function (CIA) binds and activates estrogen receptors and some orphan nuclear receptors (Sauvé et al., 2001). Interestingly, both proteins can also act as transcriptional repressors (Sauvé et al., 2001; Leong et al., 2004). Therefore, the role of the interaction of MeCP2 with these proteins and their involvement in the

reputed MeCP2 function as transcriptional activation is difficult to predict without further studies.

A number of proteins identified by us seem to be involved in regulation of pre-mRNA splicing. MeCP2 has been shown to affect the splicing profiles of various genes (Young et al., 2005). Although the protein YB-1 suggested in that study to mediate the role of MeCP2 in splicing was not detected in our purifications, it could be that the other binding partners of MeCP2 can serve the same purpose. The proteins involved in splicing and reported to interact with the putative WW-domain binding region on MeCP2 (FBP11 and HYPC; Bedford et al., 1997; Buschdorf and Strätling, 2004) were also absent from our list. However, we identified several possible candidates, such as two subunits of transformer-2 protein homolog, α (TRA2A) and β (TRA2B). Both proteins are factors promoting alternative splicing (Tacke et al., 1998; Sakashita et al., 2004). Another protein identified by us with a potential function in mRNA splicing is WW domain-binding protein 11 (WBP11), also known as splicing factor that interacts with PQBP-1 and PP1 (SIPP1) (Llorian et al., 2004; Tapia et al., 2010). Although the preliminary test by Western blotting did not confirm its interaction with MeCP2, it has to be confirmed with additional antibodies. Also, an intron-binding protein aquarius (Aqr); eukaryotic initiation factor 4A-III (Eif4a3); Pinnin (Pnn); isoform 1 of the pre-mRNA-processing factor 19 (Prpf19) and isoform 2 of RNA-binding protein Raly were on the list of MeCP2 interacting proteins and are implicated to be involved in splicing. The information about all of the proteins mentioned above were retrieved from the GeneCards database (www.genecards.org; Rebhan et al., 1997, 1998).

The results presented in this chapter provide a solid foundation for future research. The already confirmed interactions should be followed and investigated further. One of the first steps to additionally validate the importance of the detected interactions should be immunoprecipitation from wild-type brains performed with an anti-MeCP2 antibody in order to pull-down the endogenous MeCP2. If the investigated protein partner is co-immunoprecipitated in such experiment, regardless of the GFP-tag on MeCP2, it would reinforce the significance of the binding. Further

evidence would come from reciprocal IP experiments, when the immunoprecipitation with antibody against the tested protein should co-precipitate MeCP2 if the interaction is genuine. However, these types of experiments depend largely on the efficiency and specificity of used antibodies and on the abundance of tested proteins. One way to circumvent these problems is to over-express the tagged versions of investigated proteins and use specific antibodies raised against the tag. Given that the *CamkIIα* gene has already been successfully cloned by us as a GFP-fusion protein, the reverse IP (with anti-GFP antibody and probing for MeCP2 in the pull-down) can be carried out in the near future.

It is also important to remember that the detected interactions with proteins that are a part of larger complexes can be indirect. To find out which proteins are bound directly to MeCP2, the *in vitro* binding assays with purified recombinant components of the complex should be performed.

The whole list of new protein partners of MeCP2 can be studied further in detail in order to assign new functions or new proteins involved in the already postulated roles of MeCP2. To identify which of them are important for Rett syndrome each interaction should be tested for its disruption by RTT causing mutations in MeCP2.

Chapter 4: Analysis of MeCP2-importin interactions

4.1 Introduction

The eukaryotic nucleus is separated from the cytoplasm by means of a nuclear envelope. This tight lipid structure is inlaid with nuclear pore complexes, which allow small water-soluble molecules to diffuse freely in and out of the nucleus. It is generally considered that proteins larger than 9 nm diameter (corresponding to a molecular weight of 45-60 kDa) cannot passively cross the nuclear membrane (Paine et al., 1975; Nigg, 1997; Weis, 2003). However, there are reports showing that, depending on the protein, even macromolecules of size up to 110 kDa can diffuse to the nucleus (Wang and Brattain, 2007). Nonetheless, to facilitate fast and reliable nuclear import of large proteins, sometimes against the concentration gradient, active transport systems have evolved. Canonical facilitated nuclear import relies on the presence of nuclear localisation signal (NLS) on the cargo protein. A specialized class of proteins called importins α is able to recognise and bind to these sequences. Upon binding to the NLS, importin α forms a heterodimer with importin β, which docks the whole complex to the nuclear pore and translocates to the nucleus (reviewed in: Lange et al., 2007). Inside the nucleus the complex is disassembled upon binding of RanGTP, Nup50 and CAS (cellular apoptosis susceptibility protein), cargo protein is released and importins with CAS and RanGTP are recycled back to the cytoplasm (Kutay et al., 1997; Lindsay et al., 2002). GTP hydrolysis in Ran, induced by binding of Ran-BP and Ran-GAP in the cytoplasm liberates importins for another round of transport (Bischoff and Görlich, 1997).

Importins α and β belong to a karyopherin protein family. The nomenclature of importins is quite confusing; for example the equivalent of mouse importin α3 is called importin α4 in humans and vice versa (reviewed in: Goldfarb et al., 2004). Therefore, throughout this dissertation, the unambiguous symbols KPNA3 and KPNA4 will be used to describe specific importins α.

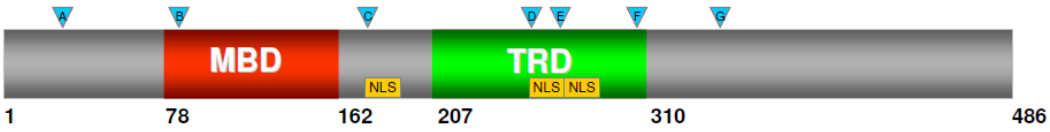
There are also other, alternative nucleocytoplasmic transport systems, which do not depend on importins nor a classical NLS. For example, calmodulin-mediated nuclear import, regulated by calcium levels in the cells, is responsible for SRY/SOX-9 delivery to the nucleus (Pruschy et al., 1994; Argentaro et al., 2003). Proteins without a classical NLS can also translocate to the nucleoplasm by binding to other proteins with a target sequence (so-called piggy-back mechanism) or by interacting directly with nucleoporins in the nuclear pore complexes (reviewed in: Wagstaff and Jans, 2009).

The interaction between MeCP2 and KPNA3 and KPNA4 identified by us is interesting not only because it can explain how MeCP2 is transported to the nucleus, but also for the reason that importins are believed to play a role in many important cell functions. For instance they can influence mitosis, nuclear envelope formation and animal development (Mason et al., 2003). In neurons importin-mediated transport is crucial for proper axon guidance, connections and plasticity (Otis et al., 2006). All these processes, when disrupted, can potentially lie in the basis of Rett syndrome. Moreover, recent report showed that KPNA2 (another member of importin α family) can also regulate nuclear size in *Xenopus* (Levy and Heald, 2010). It is established that neuronal nuclei in the brains of mouse model of Rett syndrome are significantly smaller than their wild-type counterparts (Chen et al., 2001; Yazdani et al., 2012; S. Cobb, personal communication). Therefore, we hypothesised that the strong interaction between MeCP2 and importins α could be crucial for neuronal function and development and may be connected to the nuclear size phenotype.
I decided to investigate closely a possible relationship between these proteins and Rett syndrome. The detailed mapping of the importin interaction sites on MeCP2 was followed by studies of the influence of *Mecp2* knock-out on the distribution and abundance of importins in mouse brains.

4.2 Mapping the MeCP2-importins interactions

As mentioned above, importins α are known to bind to nuclear localisation signals (NLS). These are characterised by a stretch of basic amino acids (aa) and can be either monopartite (7 aa) or bipartite (17 aa), where basic aa are separated by a

linker (Robbins et al., 1991). To identify the potential binding sites on MeCP2, its sequence was analysed bioinformatically using PSORT II algorithm (Nakai and Horton, 1999). It revealed ten putative NLS sequences (7 monopartite and 3 bipartite, Fig. 4.2.1) that could be bound by importins.



Type of NLS	Used name	Sequence	Start aminoacid
Monopartite ▼	A	PLKFKKV	25
	B	PKQRRSI	81
	C	PPKKPKS	172
	D	PGRKRKA	251
	E	PKKRGRK	265
	F	PIKKRKT	302
	G	PGRKSKE	342
Bipartite NLS	NLS1	KKPKSPKAPGTGRGRGR	174
	NLS2	RKRKAEADPQAIPKKRG	253
	NLS3	RKPGSVVAAAAAEAKKK	270

Figure 4.2.1 Prediction of Nuclear Localisation Signals in MeCP2 sequence. PSORTII algorithm was employed to predict monopartite and bipartite NLSs in MeCP2_e2 human sequence.

To test which of the predicted nuclear localisation signals bind importins α, a series of MeCP2 truncation mutants spanning different NLS sequences was used to map the interactions. Both KPNA3 and KPNA4 could be detected by Western blot after over-expression of full length FLAG-MeCP2 in HeLa cells followed by FLAG immunoprecipitation, confirming the interaction and the feasibility of the method (Fig. 4.2.2). The detailed binding analysis showed that MeCP2 truncations containing NLS2 interact with both importins (such as fragments 202-486 and 1-273; Fig. 4.2.2A and black bars in Fig. 4.2.2B). Truncations missing or disrupting NLS2 (for example 269-486, 1-262) failed to bind tested importins (Fig. 4.2.2A and white bars in Fig. 4.2.2B). Therefore, another MeCP2 mutant missing only that sequence (MeCP2 ΔNLS2) was produced and it also failed to bind KPNA3 and 4 (Fig. 4.2.2).

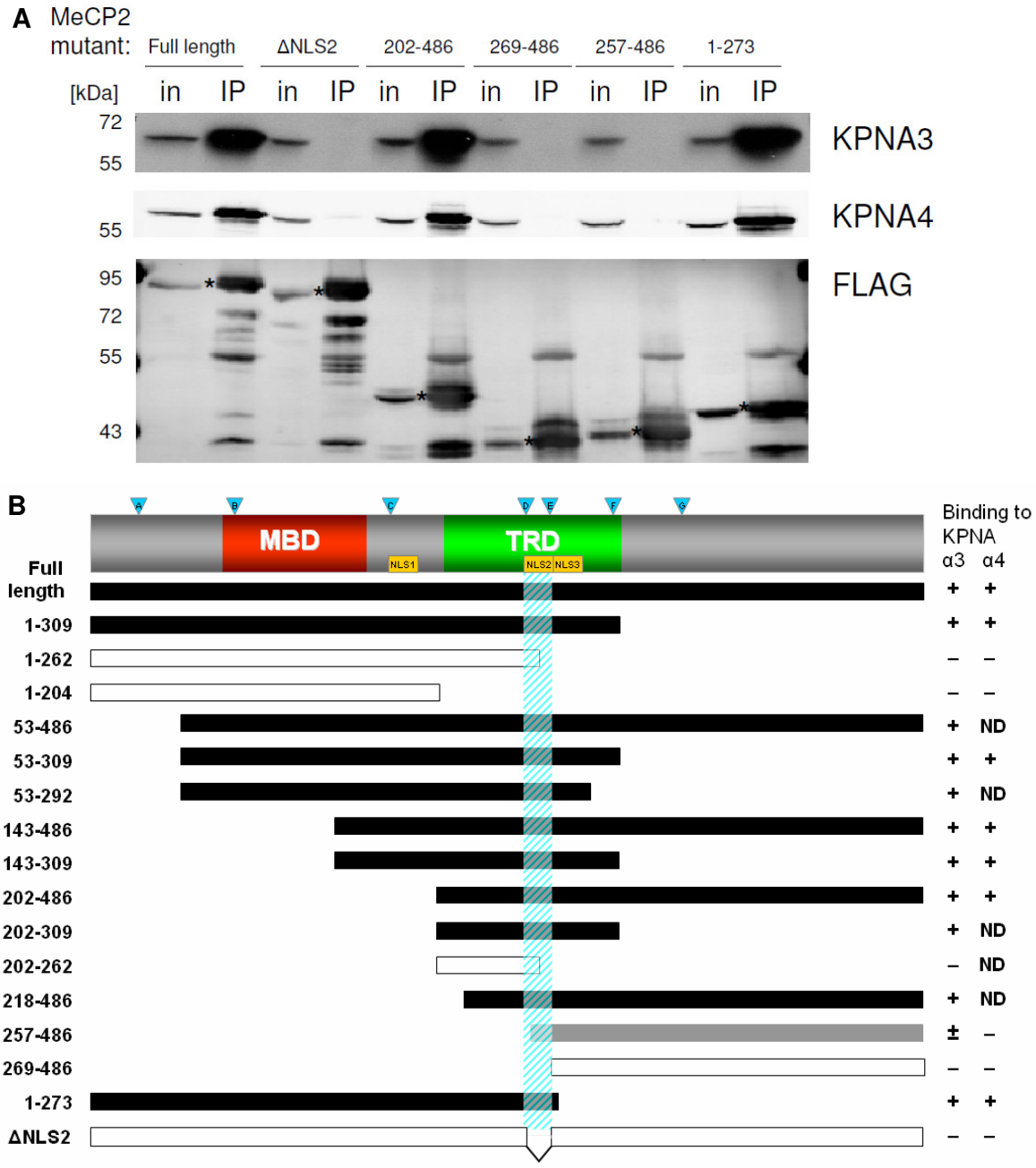


Figure 4.2.2. Mapping of MeCP2-importins interactions. (A) Western blot showing interactions with KPNA3 and KPNA4 with full-length and some of the tested MeCP2 truncations. in – input, IP – immunoprecipitation. Re-probing the membrane with anti-FLAG antibody confirmed equal expression and immunoprecipitation of MeCP2 truncations. Asterisks between inputs and IPs mark the specific band for each of them. (B) A cartoon summarising the interactions with all of the tested MeCP2 truncations. Black bars represent fragments bound by tested importins (+); white colour indicates lack of the binding (–) and grey bar stands for intermediate binding in one of two biological replicates of the experiment (±). ND – binding of KPNA4 to this MeCP2 truncation was not determined.

A very weak KPNA4 band in the ΔNLS IP was not observed in two repeats of this experiment. These results refined mapping of the interactions to a 253-269 region of MeCP2 (Fig. 4.2.2B). All of the MeCP2 truncation mutants were expressed and immunoprecipitated at comparable levels as assessed by re-probing the Western blot membrane with anti-FLAG antibody (Fig. 4.2.2A, asterisks).

4.3 Role of importins in nuclear transport of MeCP2

The NLS confirmed to interact with importins α is the same sequence reported previously to be a functional NLS of MeCP2 (Nan et al., 1996). MeCP2 fragments missing this NLS (1-261, 1-178) were localised mostly in the cytoplasm, with some 'leakage' detected to the nucleus. However, this study was performed using large fusion of MeCP2 fragments with β-galactosidase, which is known to form a homotetramer of total weight exceeding 1.8 MDa (Jacobson et al., 1994) whose size could affect the localisation of the tested constructs. Others have reported that there must be another functional NLS in MeCP2, because MeCP2 [168X] fused with GFP was detected in the nucleus (Kumar et al., 2008), although it does not have any of the predicted NLS sequences. In light of our data that MeCP2 interacts with importins only at NLS2 I decided to test myself the localisation of full-length MeCP2 and its mutant truncated at the position 255 (MeCP2 [255X]), hence lacking the NLS2. NIH-3T3 mouse fibroblast cells were transfected with plasmids expressing wild-type or mutant MeCP2 fused with GFP under the CMV promoter. The cells were fixed, immunofluorescently stained for MeCP2 and observed under the confocal microscope. In 54 out of 55 tested cells full-length protein was localised in the nucleus in the characteristic DAPI bright foci (Fig. 4.3A), which are known to be regions of methylated pericentromeric heterochromatin (Miller et al., 1974). On the contrary, in 56% of 69 examined cells transfected with GFP-MeCP2 [255X] mutant, the protein was detected in the nucleus as well as in the cytoplasm. This was more prominent with the red antibody signal than in the GFP channel (Fig. 4.3B). However, in 44% of cells the localisation of truncated MeCP2 was identical with full length protein, being detected solely in the nucleus (Fig. 4.3C). As imaging of cells fixed with paraformaldehyde sometimes can lead to artefacts (Schmiedeberg et al., 2009), the cells transfected in the same way were also observed live under the

fluorescence microscope. The wild-type protein localised again in the nucleus in the characteristic bright foci. Localisation of GFP-MeCP2 [255X] was even more similar to full length protein with approximately 54% cells with green GFP signal solely in the nucleus and less than a half of cells with detectable MeCP2 in the cytoplasm (data not shown). These experiments confirmed previous observations that MeCP2 is able to localise in the nucleus even without the NLS2, although its distribution is slightly impaired. The possible reasons for it, proposed ways to test them and potential consequences are discussed later.

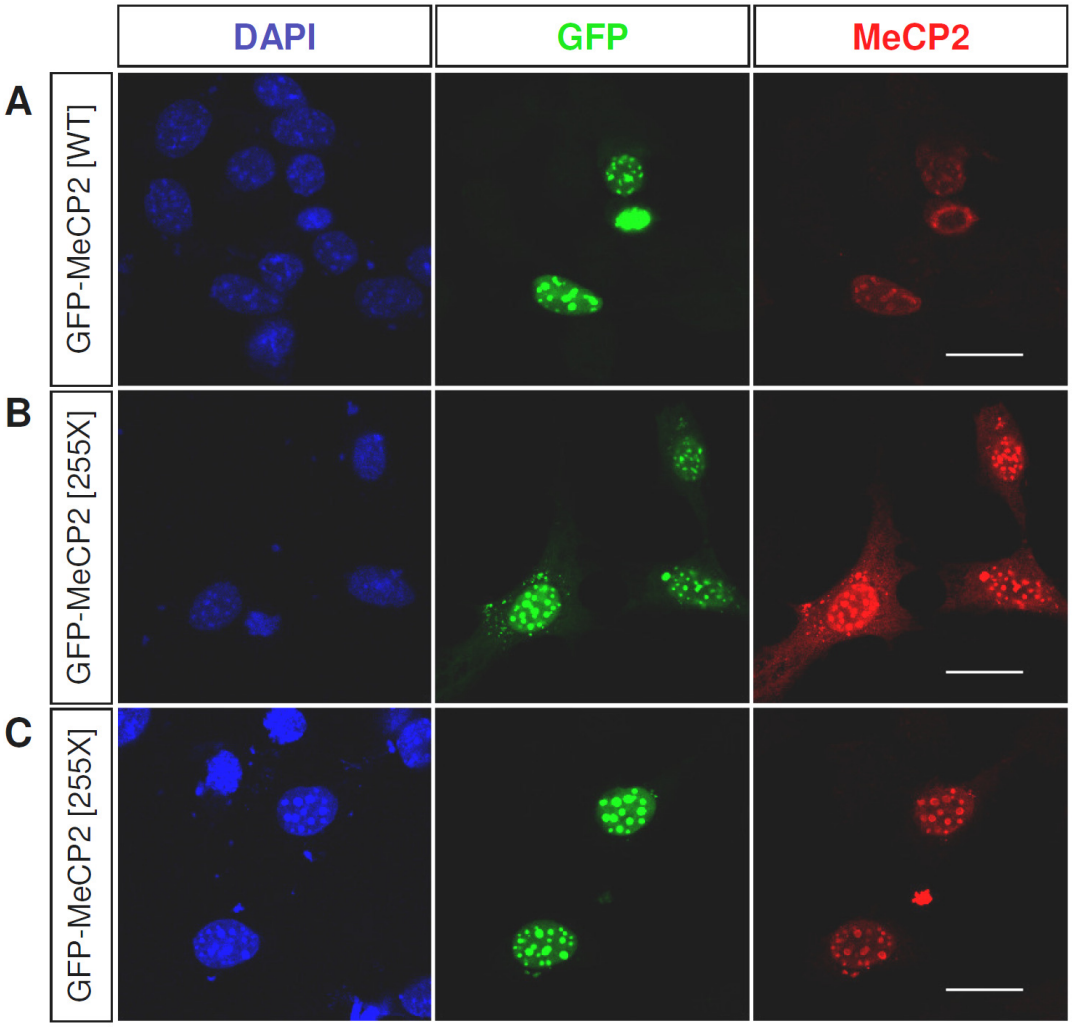


Figure 4.3. MeCP2-GFP localises in the nucleus even when truncated at position 255. NIH-3T3 cells were transfected with wild-type GFP-MeCP2 or GFP-MeCP2 [255X] and images were taken under the confocal microscope. (A) Full length MeCP2 exhibited only the nuclear localisation. (B) In 56% of cells transfected with truncated MeCP2 it localised in the cytoplasm as well as in the nucleus. (C) In 44% of cells MeCP2 [255X] was observed only in the nucleus, similarly to full length protein. The DNA constructs were gifts from L. Schmiedeberg. Scale bar: 20 µm.

4.4 Effect of MeCP2 knock-out on importins distribution

4.4.1 Analysis of importins distribution in Mecp2-GFP$^{+/-}$ brain

Importins α are known to shuttle between cytoplasm and nucleus while transporting their cargo. The strong interaction between them and MeCP2 in the nucleus (40% of total KPNA3 and ~93% of KPNA4 are bound by MeCP2; Chapter 3.3) led us to the idea that importins can be retained in the nucleus in wild-type neurons and that the release of this retention in *Mecp2* knock-out brains may lead to defects that contribute to Rett syndrome. To test this hypothesis I looked at the distribution of importins α in wild-type neurons and in cells lacking MeCP2. The direct comparison was possible by immunohistochemistry in the brain slices of a female heterozygous mouse *Mecp2-GFP*$^{+/-}$. In these animals one allele of MeCP2 has a fusion GFP protein knocked-in under the endogenous *Mecp2* promoter and the *Mecp2* gene on the other allele is deleted. MeCP2-GFP is expressed in only half of the neurons, which are mixed with the *Mecp2* knock-out neurons in the brain due to random X chromosome inactivation. The presence of MeCP2 can be visualized by the endogenous GFP signal and the superimposed picture of immunofluorescent staining of protein of interest should show if lack of MeCP2 has any influence on the localisation of the tested protein. To confirm the feasibility of the method the brain slices were first immunostained with antibody raised against MeCP2. Confocal microscopy scans were carried out in the CA1 region of the hippocampus and in the dentate gyrus, which are both known to be enriched in neurons (Buckmaster and Schwartzkroin, 1995). Green signal from MeCP2-GFP fusion protein was detected in approximately 50% of the nuclei, which were visualised by staining with DAPI (Fig. 4.4.1-1). Co-staining with anti-MeCP2 antibody gave a very similar picture and the great overlap between red and green signals allowed me to consider each green signal in the subsequent experiments as MeCP2 positive (Fig 4.4.1-1). Brain slices containing the hippocampus were stained with two different antibodies raised against KPNA4. Although the quality of staining was not as good as in the case of MeCP2, the importin distribution has been preliminarily assessed. Analysis of 60 nuclei on 3 brain slices suggested that the localisation of KPNA4 was not affected by the

presence or lack of MeCP2 (Fig. 4.4.1-2). KPNA4 seemed to be nuclear and not concentrated in the heterochromatic foci, regardless of the presence or absence of MeCP2.

Immunofluorescent staining with antibody raised against KPNA3 was performed in a similar fashion, however the low signal/noise ratio made the analysis impossible (data not shown).

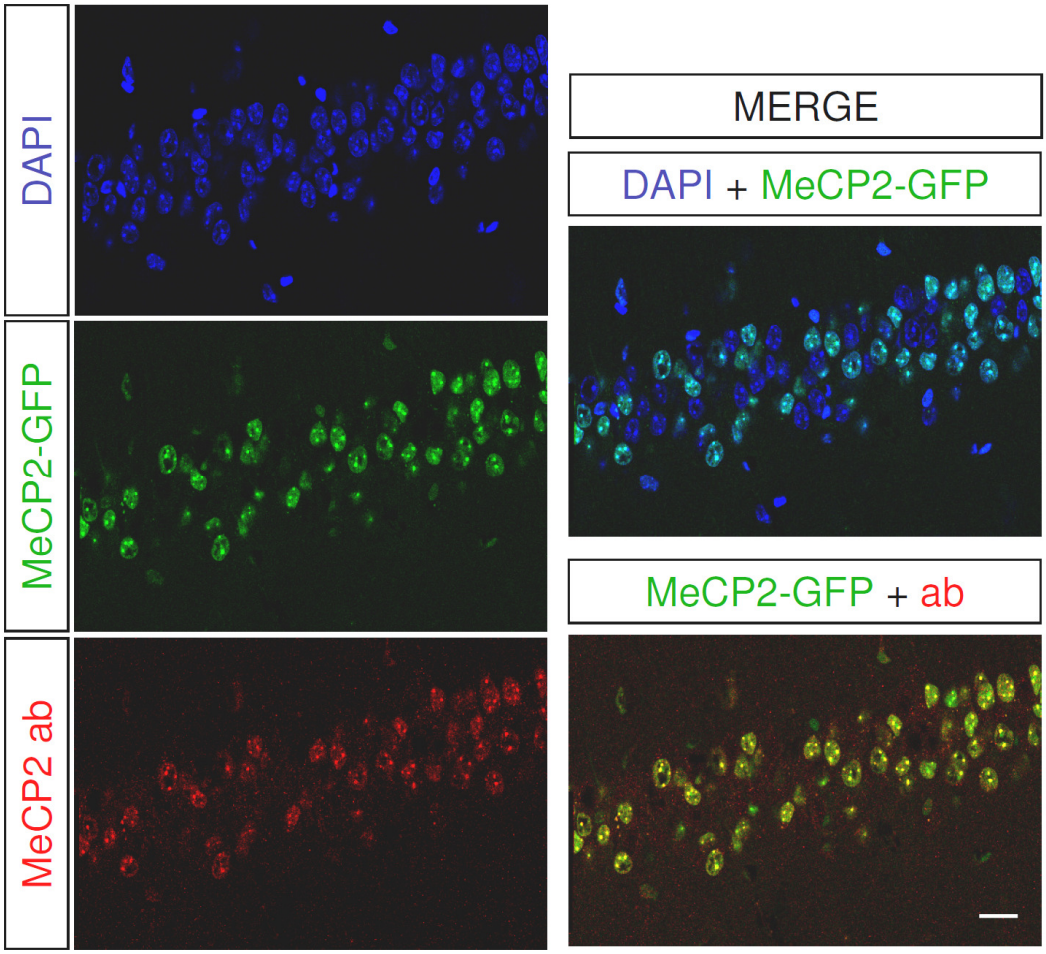


Figure 4.4.1-1. Distribution of MeCP2-GFP in *Mecp2-GFP*$^{+/-}$ heterozygous female mouse brain. Brain slices from *Mecp2-GFP*$^{+/-}$ female mouse were immunostained with anti-MeCP2 antibody and DAPI. Confocal miscroscopy showed the presence of MeCP2 in only half of the neurons in characteristic DAPI-bright heterochromatic foci. The GFP signal of MeCP2-GFP fusion protein overlapped completely with the antibody staining. Scale bar: 20 µm.

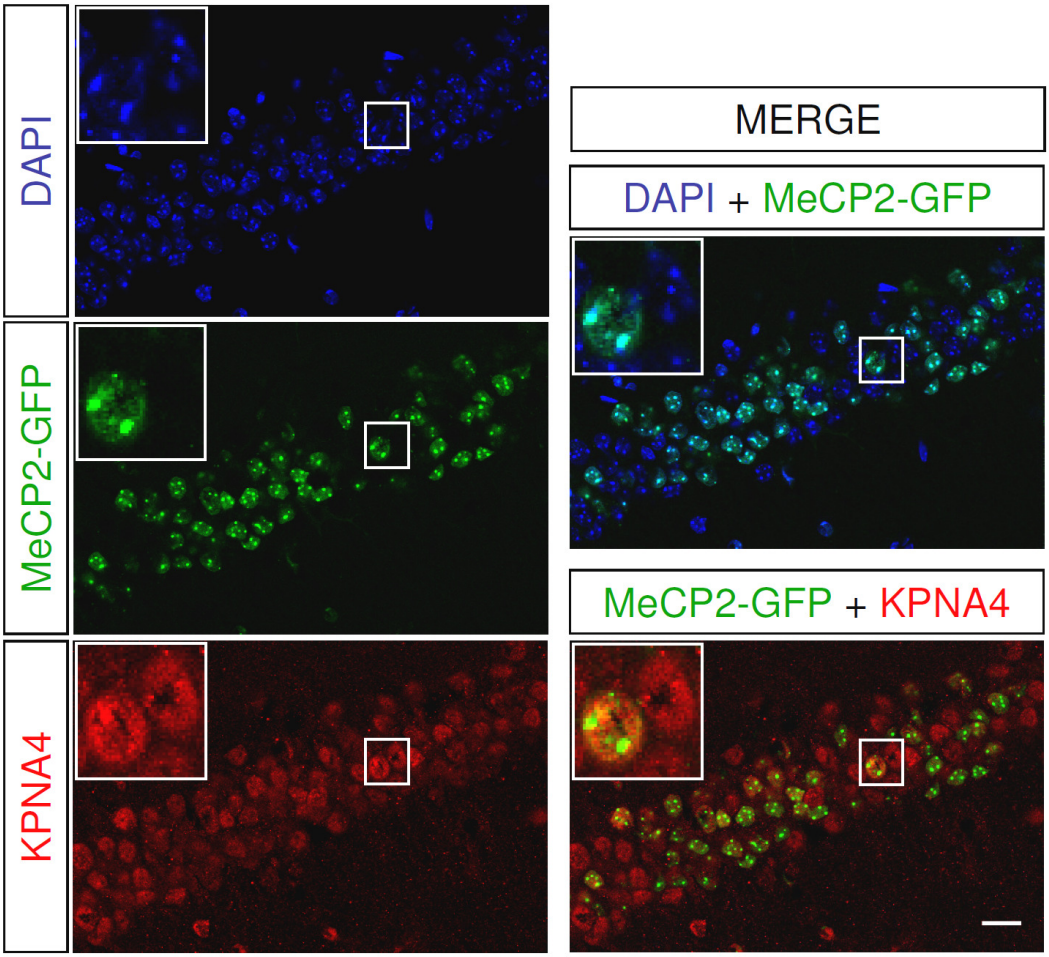


Figure 4.4.1-2. Distribution of importins in *Mecp2-GFP*$^{+/-}$ heterozygous female mouse brain. Brain slices from *Mecp2-GFP*$^{+/-}$ female mouse were immunostained with anti-KPNA4 antibody and DAPI. Confocal miscroscopy showed the presence of MeCP2 in only half of the neurons, and no difference in the distribution of KPNA4 between them. The insets are 3-fold magnifications of areas marked with white rectangles. Scale bar: 20 µm.

4.4.2 Importins distribution in biochemically fractionationated brains

Another method to test the distribution of proteins between the nucleus and cytoplasm is biochemical fractionation. The brains from wild-type and MeCP2 knock-out mice were homogenized and proteins divided into cytoplasmic and nuclear fractions. Probing the Western blot membrane with antibodies against control proteins showed their predicted distribution: GAPDH was solely in the cytoplasm,

whereas histone H2B in the nuclear fraction (Fig. 4.4.2). Additionally, as expected, MeCP2 was present only in the nuclei of wild-type brains and could not be detected in the samples from *Mecp2*-null animals. Both KPNA3 and KPNA4 proteins were detected mainly in the nuclear fraction and their distribution was not affected by the lack of MeCP2 (Fig. 4.4.2) corroborating the results obtained by immunofluorescent confocal imaging.

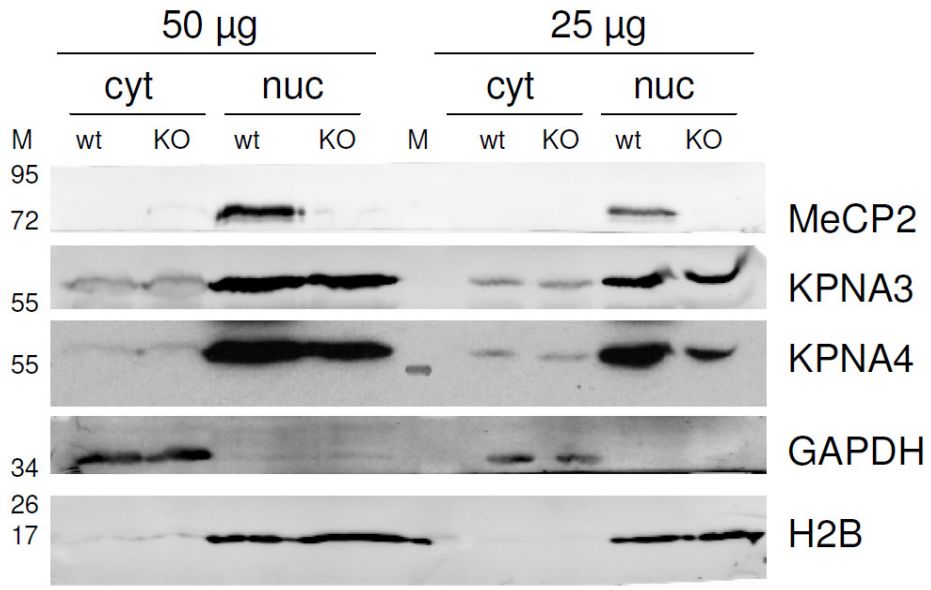


Figure 4.4.2. Lack of MeCP2 does not influence importins distribution. Protein extracts from biochemically fractionated wild-type (wt) and *Mecp2* knock-out (KO) mouse brains were separated by SDS-PAGE, Western blotted and probed with antibodies marked on the right. cyt – cytoplasmic fraction, nuc – nuclear fraction. Representative of three independent experiments.

4.4.3 The effect of *Mecp2* knock-out on mRNA levels of importins

Although lack of MeCP2 had no influence on importin distribution between cytoplasm and nucleus, we observed a weak trend towards lower levels of KPNA3 and KPNA4 proteins in *Mecp2* knock-out brains when compared to wild-type (Fig. 4.4.2 and data not shown). That led us to an alternative hypothesis that could link MeCP2-importins interactions to the Rett syndrome. In wild-type neurons the interaction with MeCP2 causes sequestration of the pool of importins α in the nucleus and leads to their increased production. In *Mecp2* knock-out brain this

hypothetical feedback mechanism is not active and in consequence importins are present at lower levels what may cause some of the Rett syndrome phenotypes. The abundance of the protein may be regulated at both transcriptional and translational levels, as well as being affected by the stability of the protein. Therefore, a comprehensive analysis of the influence of *Mecp2* knock-out on levels of both KPNA3 and KPNA4 was performed. The quantification of mRNA molecules of importins relative to *Gapdh* and assessed by RT-qPCR showed no significant effect of MeCP2 on their transcription (Fig. 4.4.3). *Kpna6* mRNA was used as a control and its expression should not depend on the presence of MeCP2 as it was not detected to interact with it. Indeed the mRNA level of *Kpna6* also stayed unchanged upon deletion of MeCP2 (Fig. 4.4.3).

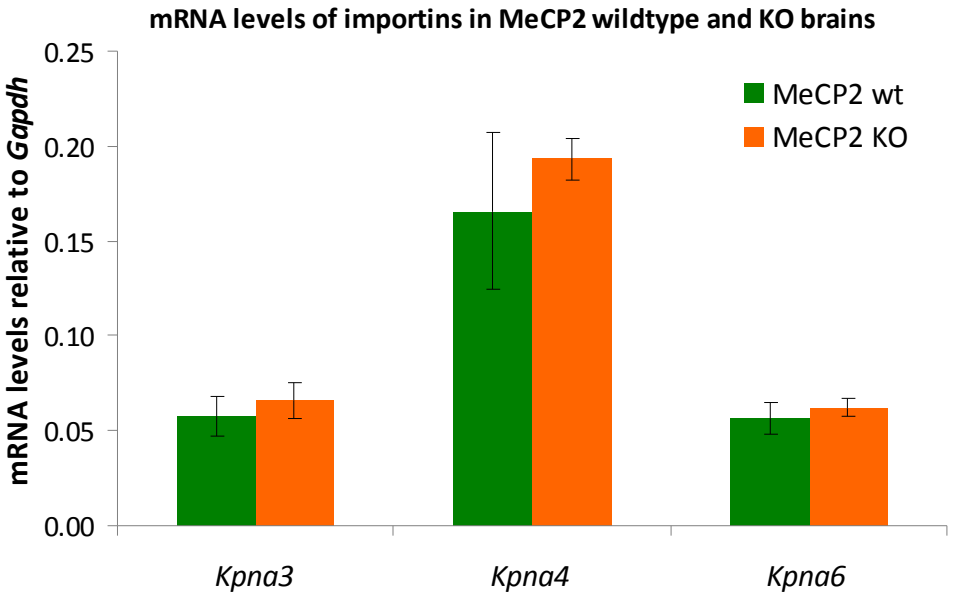


Figure 4.4.3. Knock-out of *Mecp2* does not influence the mRNA levels of importins α in the brain. qPCR analysis was performed on cDNAs from brains from two wild-type (wt) and two *Mecp2* knock-out (KO) mice (gift from Pete Skene). Error bars represent standard deviations calculated from tryplicated PCR wells.

4.4.4 The effect of *Mecp2* knock-out on protein levels of importins

The protein levels of importins were tested by quantitative Western blotting and normalized to levels of the neuronal marker NeuN. As expected, MeCP2 was

absent in the knock-out brains. No obvious difference in the KPNA3 and KPNA4 levels was detected between wild-type and *Mecp2* knock-out brains. The equal loading of the gels was additionally confirmed by even levels of total histone H3 between the samples (Fig. 4.4.4).

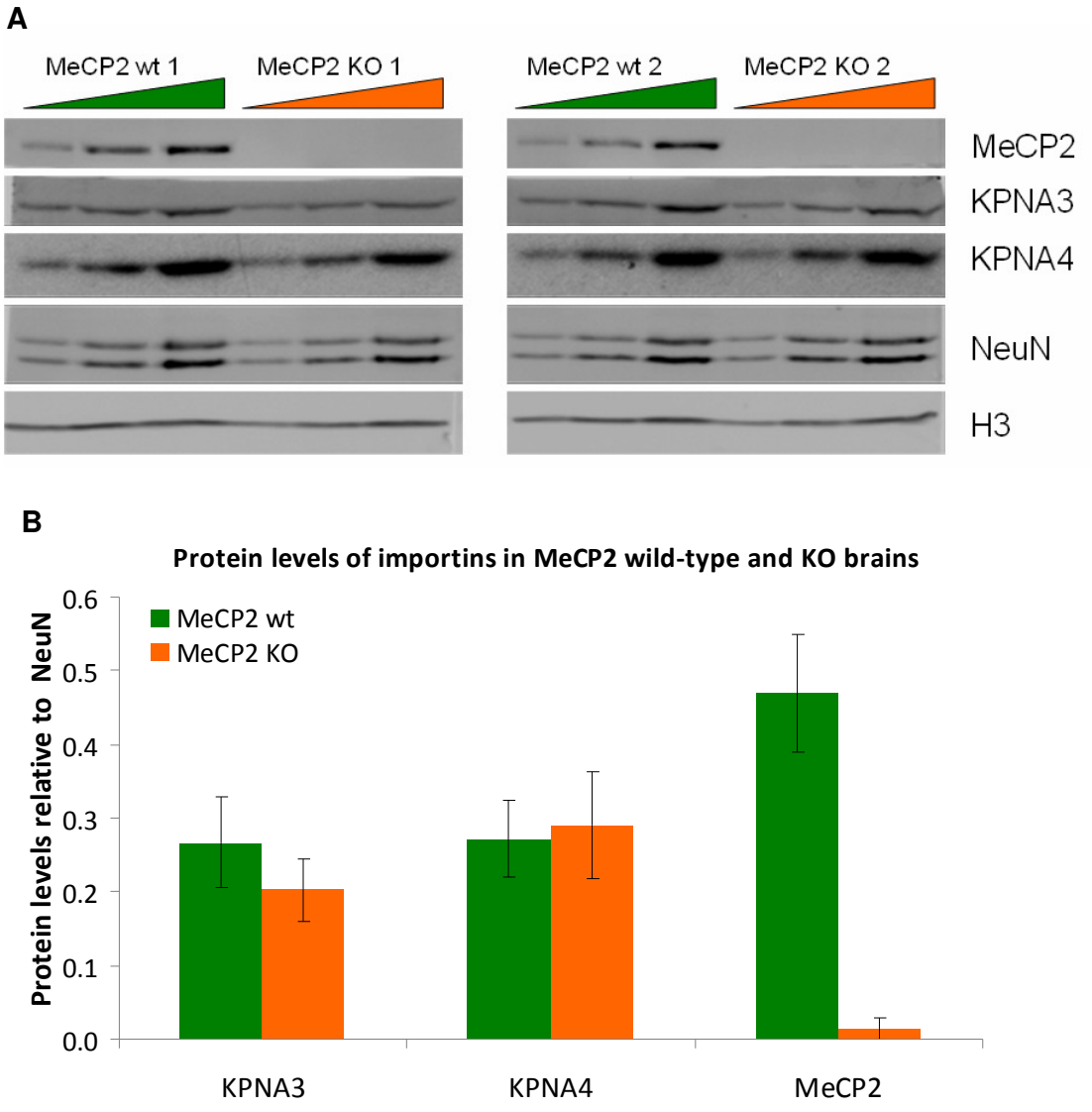


Figure 4.4.4. Knock-out of *Mecp2* does not influence the protein levels of importins α in the brain. Brains from two wild-type (wt 1 and wt 2) and two *Mecp2* knock-out (KO 1 and KO 2) mice were homogenized, proteins extracted and separated by SDS-PAGE in three concentrations increasing two-fold. (A) Western blot membranes were probed with antibodies against MeCP2, KPNA3 and KPNA4, neuronal marker NeuN and histone H3. (B) Densitometric analysis revealed no significant difference in protein levels of importins between MeCP2 wild-type and knockout brains. Representative blot of three independent experiments. Error bars represent standard deviations.

4.5 Discussion

In Chapter 3 we detected interactions between MeCP2 and mouse importins α KPNA3 and KPNA4 by mass spectrometry and validated them by Western blotting. Here we confirmed these interactions in HeLa cells, proving that they are not species, nor cell-type specific. Detailed mapping restricted both interactions to a short sequence between positions 253-269 of MeCP2, which corresponds to the previously identified functional NLS (Nan et al., 1996). The results presented here indicate for the first time that MeCP2 is transported to the nucleus via a canonical, active, importin-mediated pathway.

The fact that only one of the ten putative nuclear localisation signals in MeCP2 is interacting with identified importins α was not surprising. The algorithm used for identifying potential NLS sequences is only predictive and therefore not always accurate (Nakai and Horton, 1999). Moreover, isoforms of importin α subunit exhibit some sequence specificity (Köhler et al., 1999; Quensel et al., 2004) and do not bind to any NLS-like sequence.

MeCP2 is quite often depicted in the literature as possessing two NLS sequences (Bienvenu and Chelly, 2006; Singh et al., 2008; Marchetto et al., 2010; Gadalla et al., 2011), although deletion analysis revealed that MeCP2 localises purely in the nucleus of mammalian cells only when contains NLS2 (Nan et al., 1996). The argument for existence of another functional NLS at 173-193 region of MeCP2 comes from low resolution immunofluorescent experiments performed in *Drosophila melanogaster* SL2 cells (Kudo, 1998). Also Kumar and colleagues postulated that there is another NLS in MeCP2 since GFP-tagged MeCP2 [168X] truncation mutant can still localise in the nucleus (Kumar et al., 2008). Furthermore, in my experiments GFP-MeCP2 [255X] mutant devoid of NLS2 was detected in the nucleus. However, the biochemical data disagree with this notion as the other putative NLS sequences do not bind the tested importins and no other proteins that could be involved in the active transport were detected by mass spectrometry.

A possible explanation for the MeCP2 localisation puzzle could be that, although it is almost as big as the limit for free diffusion through the nuclear pores (60 kDa), due to its unusual elongated shape (Klose and Bird, 2004) it can get to the

nucleus by a passive mechanism. Once it is there its high affinity to methylated DNA due to the MBD domain sequesters it from the free-floating pool of MeCP2 and more molecules can diffuse to the nucleus. There is already one piece of evidence supporting this hypothesis. The localisation of short N-terminal fragment of MeCP2 (1-108) lacking NLS2 and also devoid of MBD was solely cytoplasmic (Nan et al., 1996). This free-diffusion theory would explain why the MeCP2 fragments fused to large β-galactosidase were localised mostly in the cytoplasm when lacking NLS2 (Nan et al., 1996), but much smaller fusion proteins GFP-MeCP2 [168X] (Kumar et al., 2008) and GFP-MeCP2 [255X] (this study) were detected more visibly in the nucleus. However, to convincingly test this hypothesis a series of MeCP2 constructs without NLS2 and mutated in the MBD domain, but retaining other parts of protein should be tested. A prediction of this model is that MeCP2-ΔNLS2 with disrupted ability to bind to methylated DNA will not localise in the nucleus, although this needs to be formally assessed in the future. Additionally, the localisation of MeCP2 mutants non-interacting with importins α can be tested in cells treated with 5-azacytidine or in cells with mutated *Dnmt* genes. The disruption of methylation machinery in such cells should result in lack of MeCP2 binding sites in the genome and therefore MeCP2 should not be retained in the nucleus without its functional NLS2.

Another possibility is that in the immunofluorescence experiments, expression of MeCP2 is driven by a strong CMV promoter resulting in non-physiological over-expression of MeCP2 and in consequence incorrect localisation. The fact that full-length MeCP2 was never observed in the cytoplasm and that in some cells GFP-MeCP2 [255X] was detected only in the nucleus suggests, however, that over-expression is not a problem. Further interpretational difficulties may come from the fact that tested MeCP2 truncation mutants were designed by introducing a premature stop codon in the full-length MeCP2 sequence by site-directed mutagenesis (Kumar et al., 2008; Schmiedeberg et al., 2009). If the artificially inserted stop codon occasionally is read-through by the translational machinery, it will produce a full-length protein localising in the nucleus as wild-type. That would give the wrong impression that truncated MeCP2 without NLS2 sequence is transported to the nucleus. Therefore, generation and testing truncated

gene constructs knocked-in under the endogenous promoter of MeCP2 would give more accurate answers about their localisation

The identified classical active transport system of MeCP2 may seem to be unnecessary and redundant when the protein can localise in the nucleus without it. However, an active import to the nucleus is faster and more reliable than free diffusion. In order to investigate the importance of importin mediated transport of MeCP2 it would be interesting to test how the knock-down of importin α subunits affect the MeCP2 localisation and function, for example transcriptional repression of reporter genes. This approach however may be difficult to interpret as disrupting the general nuclear import system can influence many other processes.

Conservation of fragments of protein sequence during the course of evolution suggests their functional importance. The positively charged amino acids in the 253-269 NLS2 region of MeCP2 are conserved from zebrafish to humans, implying their evolutionary advantage (Fig. 4.5; lysines and arginines are marked in red colour). Although the amino acids in the linker region of this bipartite NLS are different amongst the species, the length of the fragment is unchanged, plausibly preserving its function. On the contrary, NLS1 is shorter and NLS3 is longer in fish compared to other classes of animal kingdoms, suggesting that these sequences may be not functional (Fig. 4.5).

Interestingly, there are no registered Rett syndrome-causing mutations in the NLS2 region in the RettBASE database. This may be due to the fact that NLS sequence is quite redundant and a change of a single amino acid due to a missense mutation will not disrupt its function (Kosugi et al., 2009). Another possibility is that the database of mutations causing Rett syndrome is not complete or the mutations present in the human population have not saturated all of the possible sites for causing the disease. It is also possible that the RTT mutations on other sites of the protein affect its folding and in consequence the interaction with importins. This is easy to imagine in the case of the P225R and P322L mutations, where proline residues are substituted with more flexible amino acids that detrimentally affect the

tertiary structure of MeCP2. The effect of the most common RTT missense mutations on the binding of importins should be analysed in the future.

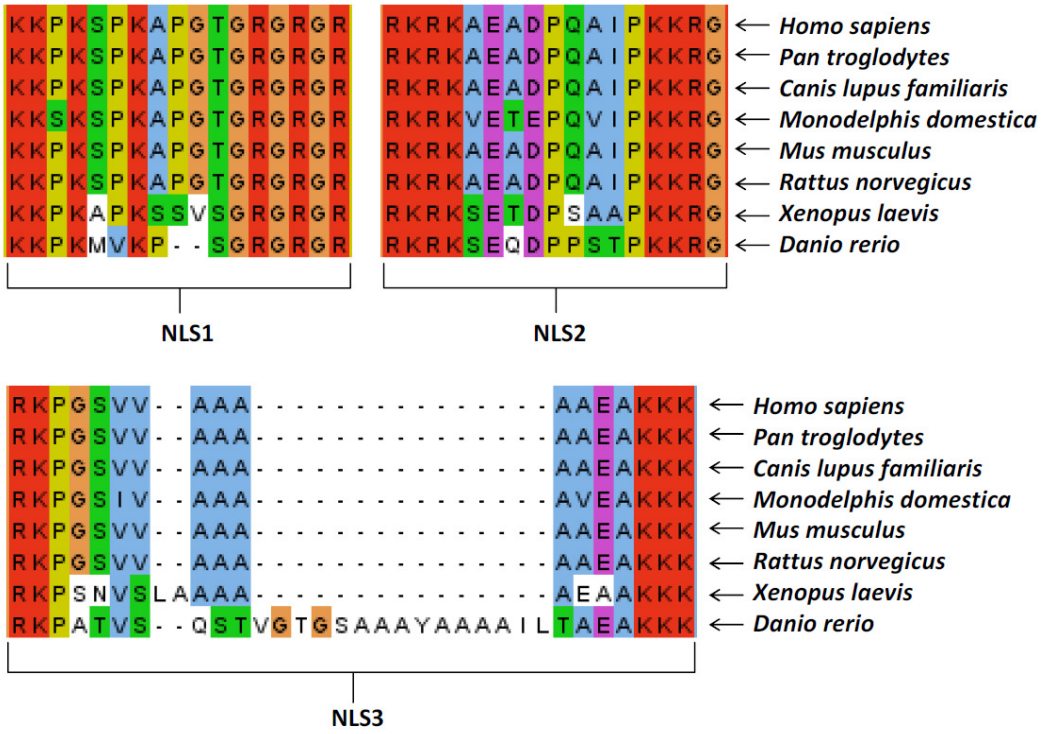


Figure 4.5. Alignment of nuclear localisation signals of MeCP2 from different animals. The sequences of putative bipartite NLS sequences of MeCP2 from human *(Homo sapiens)*, chimpanzee *(Pan troglodytes)*, dog *(Canis lupus familiaris)*, opossum *(Monodelphis domestica),* mouse *(Mus musculus)*, rat *(Rattus norvegicus)*, frog *(Xenopus laevis)* and zebrafish *(Danio rerio)* were aligned using Jalview software (Waterhouse et al., 2009).

Considering the strength of interactions between MeCP2 and importins α we hypothesised that they may have other consequences than just nuclear transport and that these functions may be related to Rett syndrome. Investigation of the effects of *Mecp2* knock-out on KPNA3 and KPNA4, however, did not reveal any disturbance of their localisation. Also, the hypothesis that the retention of importins in the nucleus by MeCP2 leads to their increased expression turned out not to be true. Total levels of KPNA3 and KPNA4 mRNA and protein were not changed in the brains of *Mecp2* knock-out mice. These findings do not mean that the relationship between these proteins has no physiological consequences, but in light of the presented evidence it seems unlikely to be a significant contributor to the Rett syndrome.

Although not of crucial significance for Rett syndrome, these findings advance our understanding of the functional domains of MeCP2. Researchers involved in studies on MeCP2 could benefit from resolving the confusion regarding the number and position of functional nuclear localisation signals. Additionally, the rapid development of sequencing techniques, their falling cost and increasing accessibility, will undoubtedly lead to a better coverage of the mutations and polymorphisms in the human population that lead to disease. If the mutations in the identified fragment of MeCP2 ever appear to be linked to any disorder, the knowledge of proteins interacting there and their role should help in developing new potential drugs and therapies.

Chapter 5: Transcriptional repression function of MeCP2 and its role in Rett syndrome

5.1 Introduction

Mass spectrometric analysis of MeCP2 interacting partners revealed binding of 5 subunits of the NCoR/SMRT co-repressor complex. This multi-subunit, megadalton complex contains either NCoR1 or SMRT in its core, and HDAC3, TBL1, TBLR1 and GPS2 associated with them via multiple interactions (Li et al., 2000; Zhang et al., 2002; Oberoi et al., 2011). It was discovered as a complex responsible for repression of transcription from unliganded nuclear receptors (Chen and Evans, 1995; Hörlein et al., 1995). However, this complex is also able to repress transcription by association with proteins other than receptors (reviewed in: Karagianni & Wong, 2007). The crucial role of the NCOR/SMRT complex in the development was shown by embryonic lethality of mouse knock-outs of its components: *Ncor, Smrt* and *Hdac3* (Jepsen et al., 2000, 2007, 2008; Bhaskara et al., 2008).

The interaction of MeCP2 with NCoR/SMRT co-repressor complex in mouse brain was validated by Western blot. Subsequently, the interaction was mapped on MeCP2 by co-immunoprecipitations of FLAG-tagged MeCP2 fragments (M. Lyst, personal communication). The complex was shown to bind to a region at the end of the TRD domain of MeCP2, between amino acids 269-309 (Fig. 5.1). Interestingly, a part of that region is a hotspot for missense mutations causing Rett syndrome. The influence of these mutations on binding of the complex was tested in the laboratory by co-immunoprecipitations. All RTT point mutations in that region: P302R, K304E, K305R and R306C, abolished the interaction (M. Lyst, personal communication). These results suggested that binding of NCoR/SMRT co-repressor complex may be crucial for the correct function of MeCP2, which is missing in RTT. As discussed above, NCoR/SMRT complex is involved in repression of transcription. Therefore, we decided to test if the mutations abolishing MeCP2-NCoR/SMRT interactions affect its repressive properties.

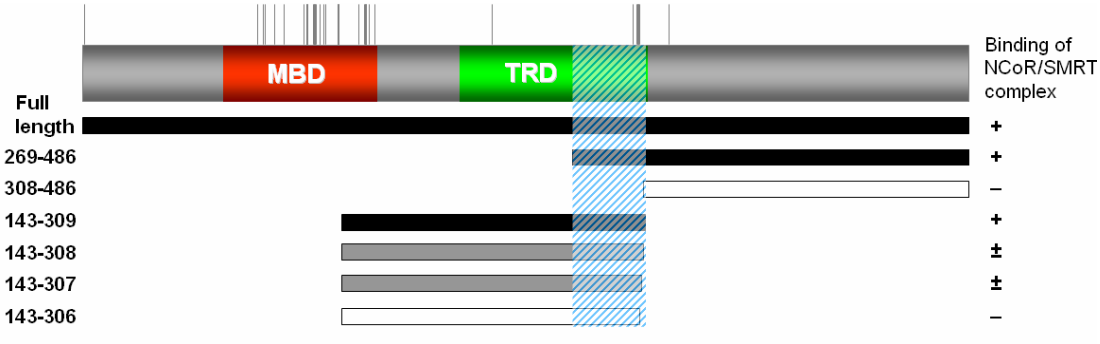


Figure 5.1. Mapping of MeCP2-NCoR/SMRT complex interactions. Summary of experiments performed by M. Lyst. Black bars represent fragments binding to the complex, grey colour indicates intermediate binding and white bars stand for lack of the binding. The mutations causing Rett syndrome are depicted as vertical lines above MeCP2.

5.2 Transcriptional repression by MeCP2 in vitro

Wild-type MeCP2 has been previously shown to repress transcription of reporter genes through its TRD domain (Nan et al., 1997, 1998). Investigation of the effects of different RTT mutations in MeCP2 on its transcriptional properties showed that nonsense mutants devoid of the NCoR/SMRT-interacting region (such as 255X and 270X) failed to exhibit repression activity (Yusufzai and Wolffe, 2000). Interestingly, MeCP2 with missense mutation R306C was reported to repress transcription as well as wild-type protein (Yusufzai and Wolffe, 2000; Drewell et al., 2002). In the light of the fact that this mutation abolishes MeCP2 association with a co-repressor complex we decided to revisit this observation. We addressed this issue by utilizing a system in mammalian cell lines where a reporter gene, *Firefly* luciferase, bearing five GAL4 binding sites was co-expressed with MeCP2 C-terminal fragment fused to GAL4 DNA binding domain (Fig. 5.2.1). A plasmid carrying *Renilla* luciferase gene was co-transfected as a control in order to be able to normalise for transfection efficiency in each sample.

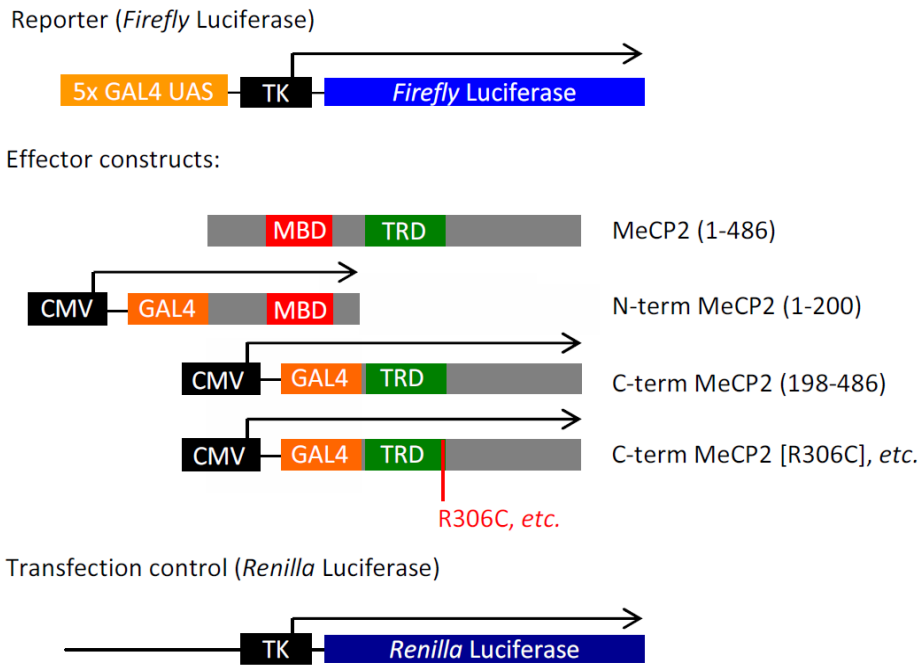


Figure 5.2.1. DNA constructs used in transcriptional repression assays. The reporter *Firefly* luciferase gene contained five GAL4 Upstream Activating Sequences (UAS) upstream of the thymidine kinase promoter (TK). Effector constructs were either N-terminal or C-terminal halves of MeCP2 fused to GAL4 DNA binding domain under the control of cytomegalovirus promoter (CMV). Point mutations were introduced in the TRD. As a transfection control *Renilla* luciferase was co-transfected with the other constructs.

The initial experiments performed in HeLa cells using equal amounts of reporter and effector plasmids showed that both C-terminal wild-type MeCP2 and its mutated version R306C are able to efficiently repress transcription. This repression was alleviated by treating cells with trichostatin A (TSA), a histone deacetylase inhibitor, demonstrating that the repression was mediated by HDACs (Fig. 5.2.2). Interestingly, the effect of TSA on MeCP2 [R306C] was much stronger than on wild-type MeCP2. That brought us to the idea that MeCP2 [R306C] may indeed be impaired in its capacity to repress transcription compared to the wild-type protein, but that this fact could be masked by its over-expression. If the system was indeed saturated with the large excess of effector protein in relation to its binding sites on the less abundant reporter gene, then even the residual repression activity of the R306C mutant could be enough to reduce its transcription. The impairment of repressive machinery of the cells by TSA treatment, however, would reduce the rate of transcriptional repression to sub-saturating levels and reveal the poorer

performance of MeCP2 [R306C] than wild-type MeCP2 in repressing the reporter gene. To test the saturation hypothesis, the amount of the effector DNA construct was titrated down to a minimal amount needed for repression by the wild-type protein. Consistent with initial experiments, at the highest concentration of 1 µg/ml of MeCP2-coding plasmid there was no difference between wild-type and R306C in their repression activity (Fig. 5.2.3, navy blue bars). However, 100-fold dilution of the effector plasmid revealed a clear difference between them. Wild-type C-terminus of MeCP2 was able to repress transcription of the luciferase reporter 6-fold when compared to luciferase activity in cells transfected with GFP instead of MeCP2, whereas the R306C mutant behaved like no-MeCP2 control (Fig. 5.2.3, light blue bars). Further 100-fold dilution of the effector plasmids resulted in no repression activity exhibited by both wild-type and mutated C-terminal fragment of MeCP2 (Fig. 5.2.3, turquoise bars). From this point 10 ng/ml of MeCP2 plasmid was used in all further experiments.

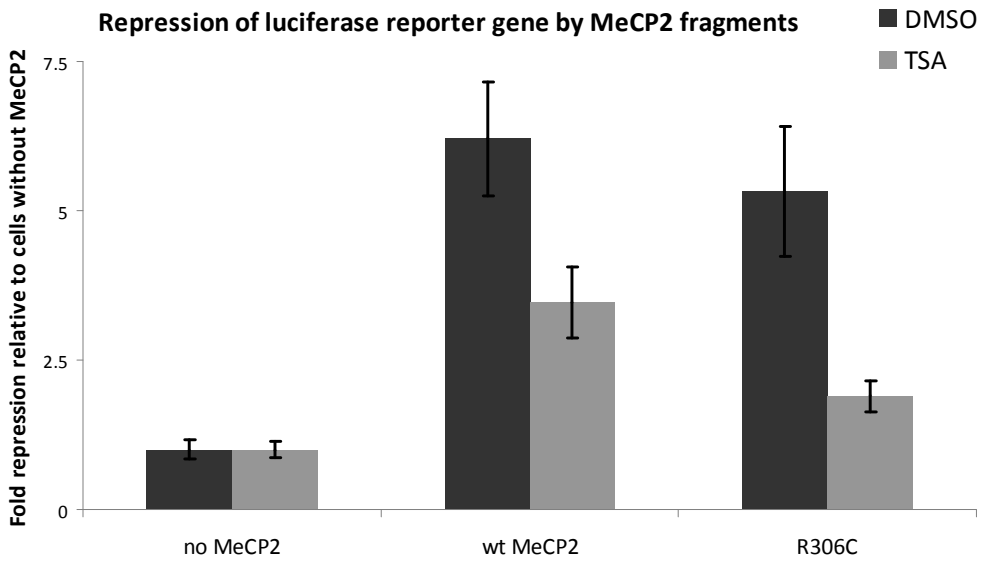


Figure 5.2.2. Over-expressed C-terminal part of MeCP2 can repress transcription of a reporter gene regardless of its R306C mutation. 1 µg of plasmid coding wild-type C-terminus of MeCP2 (wt) or its mutated version R306C were co-transfected with reporter *Firefly* luciferase gene and a control *Renilla* luciferase gene. The activities of enzymes were measured and normalised. Fold repression was calculated by dividing the obtained values by the luciferase activity in control cells transfected with GFP plasmid instead of MeCP2 (no MeCP2). The error bars represent propagated standard deviations of 4 technical replicates of each transfection.

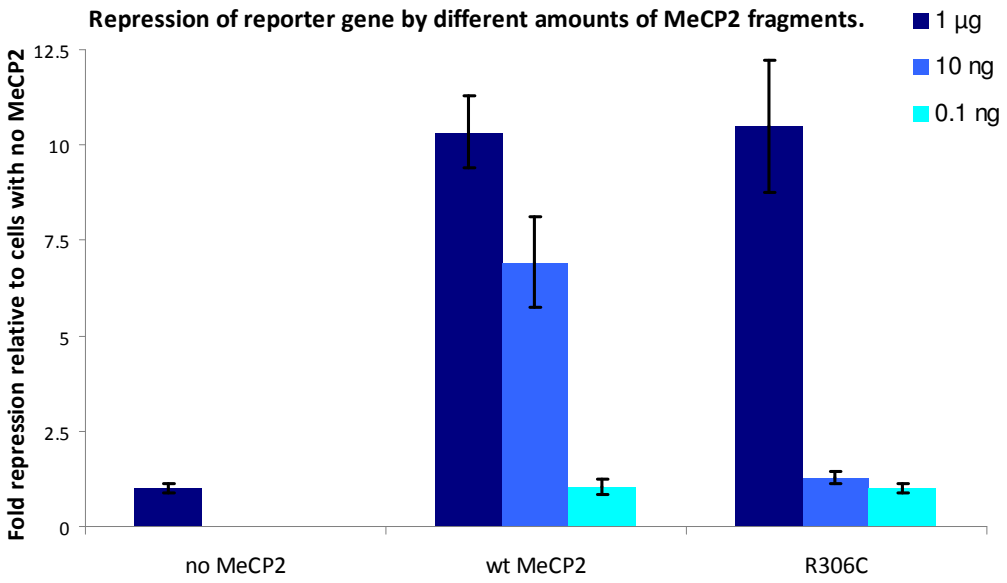


Figure 5.2.3. Reducing the amount of transfected MeCP2 construct revealed the failure in repression activity of R306C mutant. 1 µg, 10 ng or 0.1 ng of each MeCP2-coding plasmids were co-transfected with reporter *Firefly* luciferase gene and a control *Renilla* luciferase gene. The enzyme activities were measured and normalised. Fold repression was calculated as in Figure 5.2.2. The error bars represent propagated standard deviations of 3 technical replicates of each transfection.

Having found the optimal conditions to assess the repressor activity of MeCP2, different RTT missense mutations were tested. Each experiment was normalised to luciferase activity in cells transfected with GFP coding plasmid instead of MeCP2. Repeatedly, wild-type C-terminus of MeCP2 was an effective repressor, whereas R306C as well as other RTT mutants in the NCoR/SMRT-interacting region (P302R, K304E and K305E) failed to repress transcription (Fig. 5.2.4). The N-terminal half of MeCP2 was also tested as a control and, as expected, showed no repression. The differences between fold repression by wild-type C-terminal part of MeCP2 and each of the tested mutants were statistically significant (unpaired two-sample Student's t-test assuming unequal variances; $p < 0.05$). Furthermore, the repressive ability of the C-terminal fragment of MeCP2 mutated outside of the 269-309 region was assessed. P225R repressed transcription as well as the wild-type protein, which agrees with the observation that it is able to bind the NCoR/SMRT co-repressor complex. Cell treatment with TSA alleviated the repression and had no effect on the already non-repressing mutants (Fig. 5.2.4). These experiments were performed in NIH-3T3 cells as they survived TSA treatment for 48 hours better than HeLa cells.

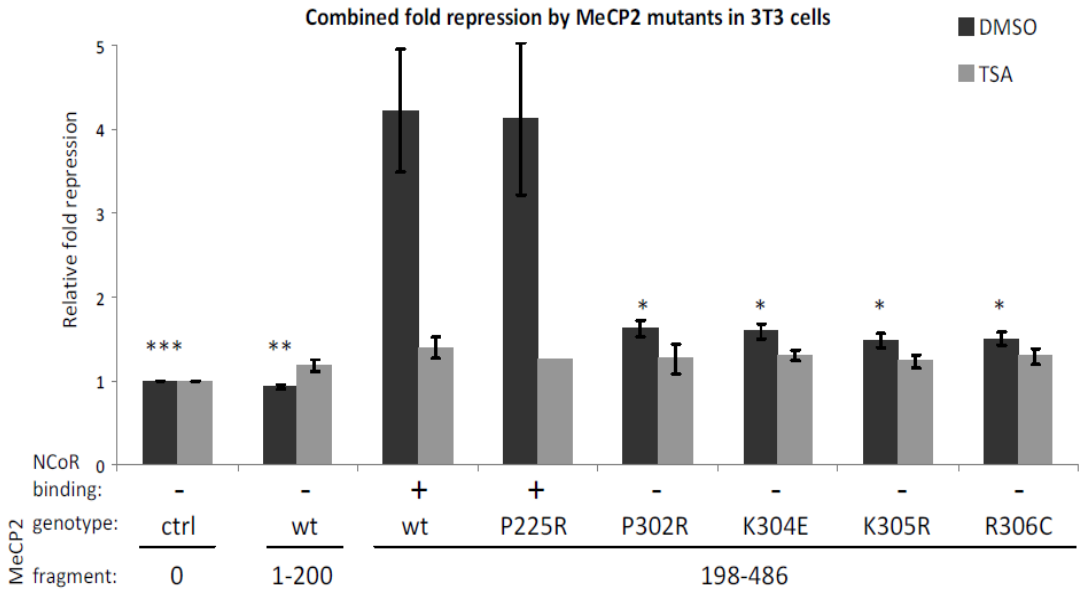


Figure 5.2.4. MeCP2 missense mutations in the NCoR/SMRT-interacting region fail to repress transcription in NIH-3T3 cells. 10 ng of plasmid coding each of the tested MeCP2 mutant was co-transfected with luciferase genes and measured and normalised as in Figure 5.2.3. wt – wild-type, ctrl – control cells transfected with GFP plasmid instead of MeCP2. The error bars represent standard errors of the means of 2-6 independent biological experiments. Asterisks above the bars represent the p value of two-tailed t-test when compared to wild-type C-terminal half of MeCP2. * - $p<0.05$, ** - $p<0.01$, *** - $p<0.001$. Data of binding of each MeCP2 mutant to NCoR/SMRT complex was obtained from M. Lyst.

Additional DNA constructs with mutated GAL4-MeCP2 were designed in order to investigate the effects of further mutations in the NCoR/SMRT-interacting region on transcriptional repression mediated by MeCP2. New tested mutants included two less common mutations causing Rett syndrome (K305E, R306H) and three mutants with amino acid changes not reported as giving rise to RTT (L301P, I303T, K307E). Another tested mutation outside of the RTT cluster, A279V, was reported only once as causative to the syndrome in Korean patients and without testing the genotype of the parents of the patients (Chae et al., 2004). Analysis of binding of all these mutants to NCoR/SMRT complex by co-immunoprecipitations revealed a lack of interaction in case of mutations between positions 301-306, intermediate binding to K307E and wild-type-like binding to A279V (M. Lyst, personal communication). The luciferase repression assay correlated well with the binding data, showing lack of repression by MeCP2 mutated between 301-306,

wild-type-like repression by A279V and intermediate repression by K307E mutant (Fig. 5.2.5). This set of experiments was performed in HeLa cells and normalised to cells transfected with the N-terminal part of MeCP2, therefore these results could not be directly combined with the similar experiments performed in NIH-3T3 cells (Fig. 5.2.4). However, the facts that the observed phenomenon is not only mouse-specific but also true in a human cell line and that the fold repression values are comparable, support the validity of presented discoveries

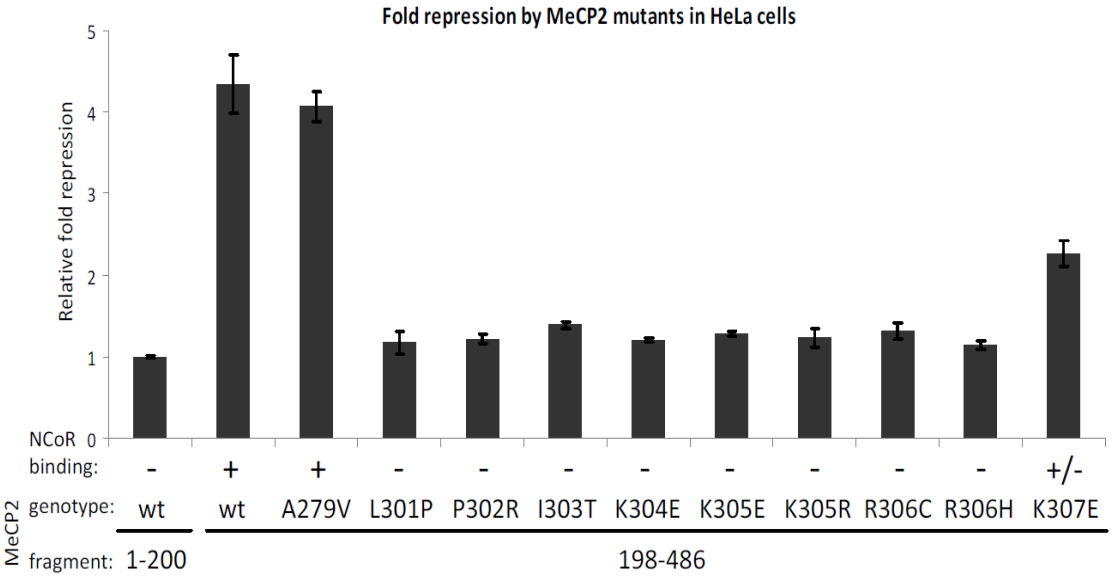


Figure 5.2.5. MeCP2 missense mutations in the NCoR/SMRT-interacting region fail to repress transcription in HeLa cells. HeLa cells were transfected as in Figure 5.2.4. Fold repression was calculated by dividing the obtained values by the luciferase activity in cells transfected with N-terminal half of MeCP2. The error bars represent propagated standard deviations of two technical repeats.

An important factor that has to be taken into account when comparing the effects of mutations in MeCP2 on its ability to repress reporter gene transcription is the expression level of each mutant. Since the amount of MeCP2 effector plasmid used in each luciferase experiment was very low it proved difficult to detect expressed proteins by conventional methods. Even the ultra-sensitive Western blot performed from HeLa cells transfected with the amount of MeCP2-coding plasmid used in the luciferase assays (10 ng) was barely able to detect the expressed protein

using the antibody against GAL4 DNA binding domain (GAL4-DBD; Fig. 5.2.6A) or against MeCP2 (data not shown). However, amongst the unspecific bands, in all lanes apart from a control cells lane (not transfected with effector plasmid) there was a visible band corresponding to predicted size of each of the fragments of MeCP2 (Fig. 5.2.6A). The levels of both wild-type N- and C-termini of MeCP2, as well as R306C, were comparable. The proteins were equally loaded on the gel as assessed by the intensity of an unspecific band at 70 kDa (an asterisk in Fig. 5.2.6A). The effect of TSA on histone acetylation levels was tested on the same Western blot and, as expected, revealed hyper-acetylation of histone H3 upon cell treatment with TSA in a dose-dependent manner (Fig. 5.2.6A). Each group of three bands corresponding to acetylated histones H3 in cells subjected to different treatments could also be used as a confirmation of even loading of the proteins on the gel.

To confirm the comparable expression of GAL4-MeCP2 constructs, Western blots were also performed on samples from NIH-3T3 cells transfected with 10-fold higher amounts of the plasmids than in the reporter gene assays. Such approach allowed easier detection of the effector proteins when probed with the antibody against the GAL4 DNA binding domain. N-terminal half of MeCP2, as well as wild-type, R306C and K304E mutants of the C-terminal half showed comparable expression when normalised to the levels of histone H3 (Fig. 5.2.6B).

These results do not definitely prove that all of the MeCP2 mutants used in the luciferase assays are expressed equally and ideally more of them should be tested in the future. However, the facts that R306C is expressed on the same level as wild-type, and that P225R and A279V mutants are able to repress transcription suggest that the other mutants could also be expressed at comparable levels. In summary, the results of reporter gene assays showing the failure of transcriptional repression by MeCP2 [R306C] and other mutants disrupting the interaction with NCoR/SMRT can be treated with confidence.

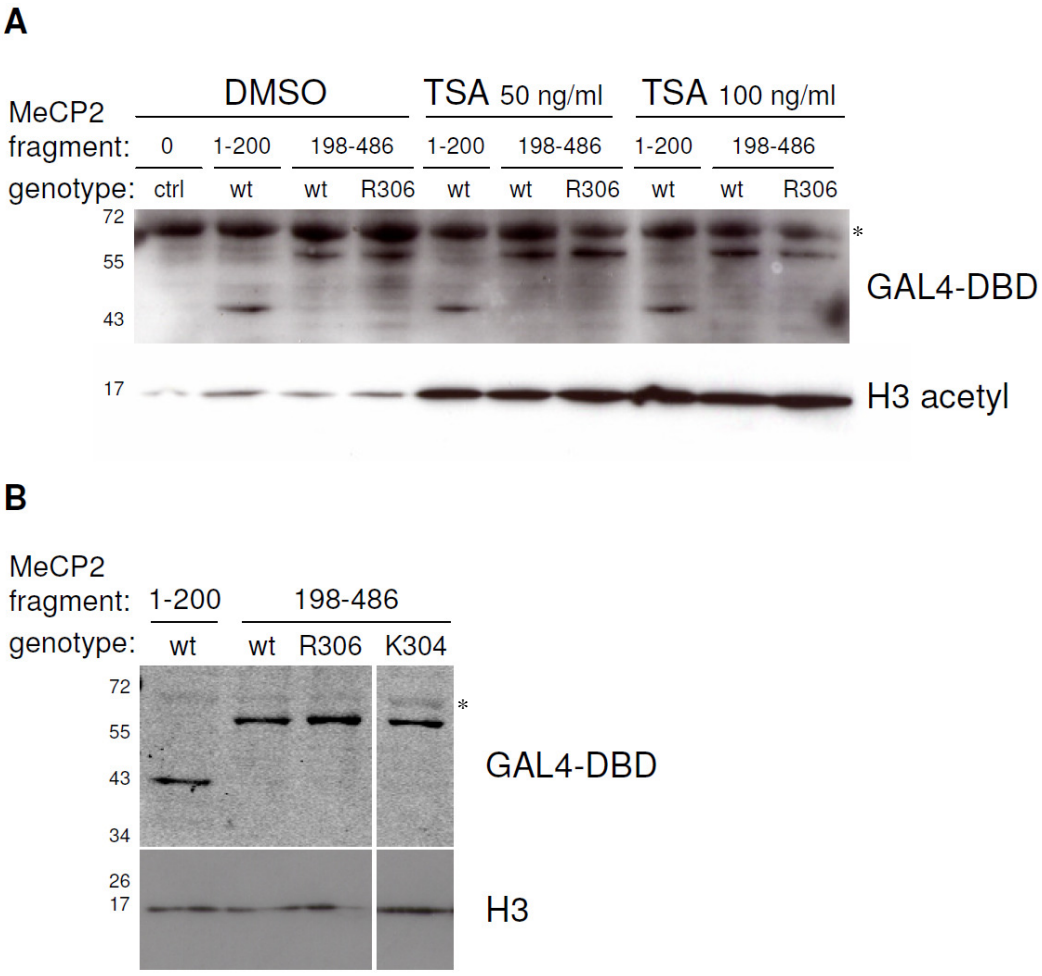


Figure 5.2.6. Equal expression of used MeCP2 mutants. (A) Ultra-sensitive Western blot showing the expression of 10 ng of MeCP2 DNA constructs in HeLa cells and the effect of TSA on the acetylation levels of histone H3. (B) Western blot showing the expression of 100 ng of MeCP2 DNA constructs in NIH-3T3 cells. An asterisk indicates the unspecific band, which can be used as a gel loading control.

5.3 Neuronal differentiation system

MeCP2 is a neuronal protein and therefore plays a role in the Rett syndrome primarily in neurons. To study the effect of RTT mutations on its properties in a more physiologically relevant system than HeLa or NIH-3T3 cell cultures, it would be the best to generate neurons with mutated MeCP2. As embryonic stem cells (ESCs) are easier to genetically alter than mature neurons, a series of ES cell lines harbouring different mutations in MeCP2 have been generated in the laboratory with a plan to produce knock-in mice (J. Guy, C. Merusi, J. Selfridge). The MeCP2-GFP knock-in ES cell line (MeCP2-GFP [WT]) has already been successfully used in making a mouse. The T158M and R306C mutations were introduced in the same locus in order to study the two most common RTT mutations disrupting two characterized domains of MeCP2 (MBD and TRD, respectively). The stem cells can be differentiated into neurons using protocol established by Bibel and co-workers (Bibel et al., 2004; Fig. 5.3.1). ESCs are first grown for 4 days on non-adhesive plates without leukaemia inhibitory factor (LIF) to produce embryoid bodies (EB) and for a further 4 days in medium containing retinoic acid (RA) to drive their differentiation into neuronal progenitors. Subsequently, the cells are harvested and seeded on laminin-coated dishes and grown in a neuronal medium. After 5-6 days in culture the cells express MeCP2 and are considered as mature and relatively pure neurons, as can be assessed by Tuj1 (neuron-specific class III beta-tubulin) staining (Fig. 5.3.1). Another neuronal marker used for detecting neurons is neuronal nuclear antigen (NeuN). Counting the ratio of NeuN positive cells over all healthy nuclei (excluding the clearly apoptotic cells) in 5 different neuronal differentiations confirmed the high purity of each culture (> 80% neurons). Staining for a glial fibrillary acidic protein (GFAP), an astrocyte cells marker, corroborated the notion that only a minority of cells in the tested cultures were glial (Fig. 5.3.2).

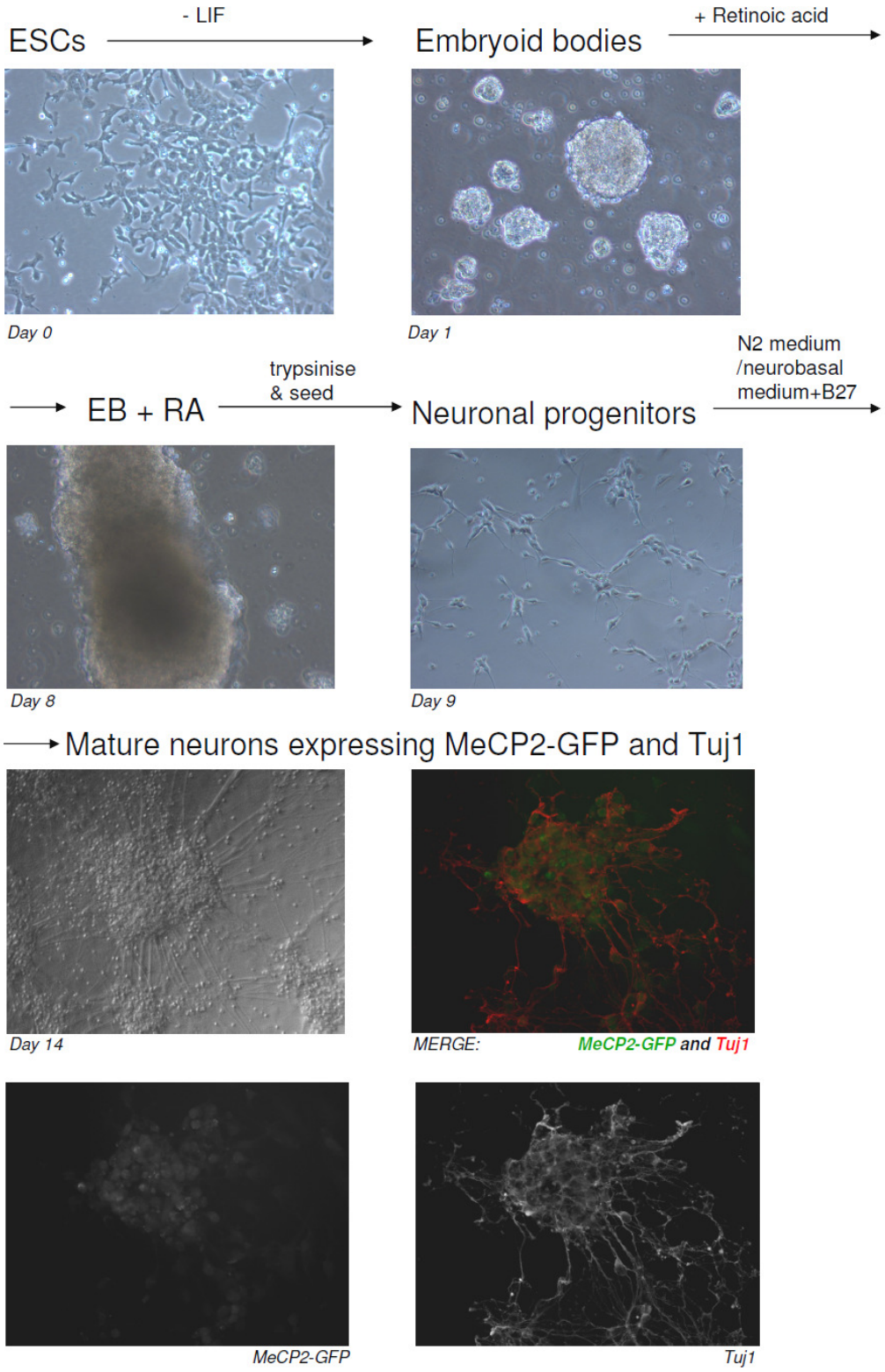


Figure 5.3.1. ES cells differentiation into neurons. The procedure of stem cells differentiation and representative pictures of milestone stages. See text for the used abbreviations.

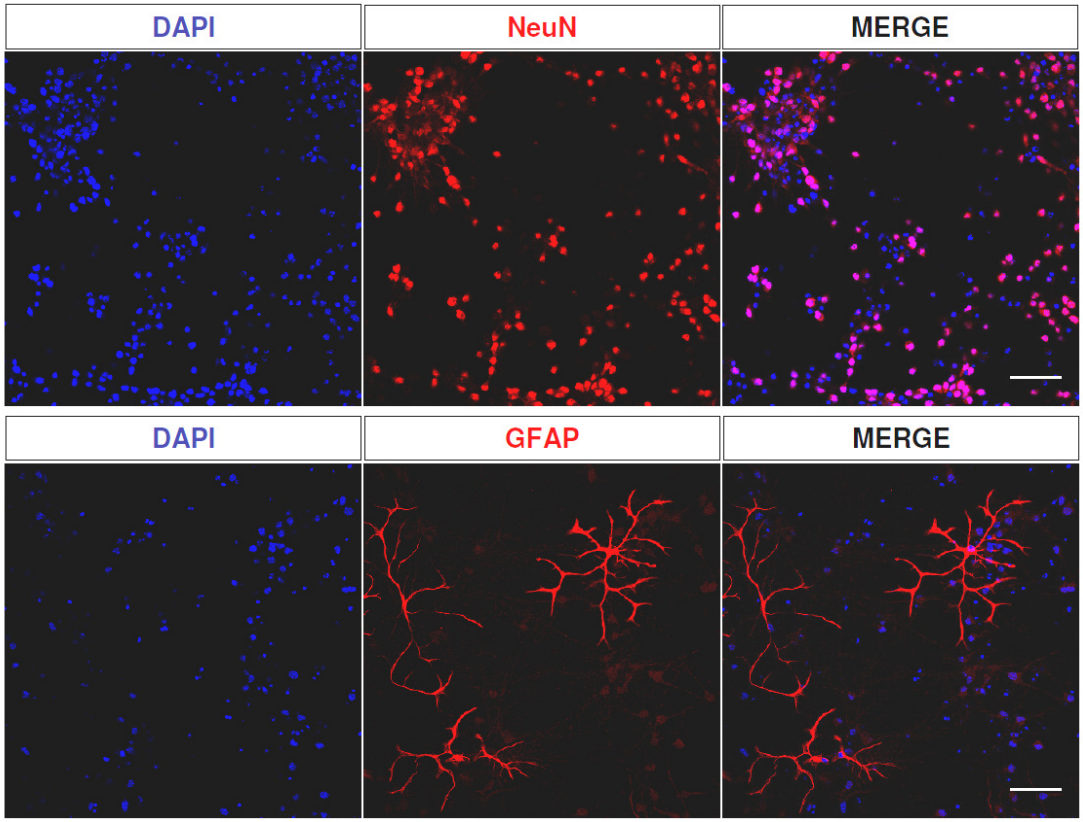


Figure 5.3.2. Purity of ES cells-derived neurons. Confocal images of neurons derived from MeCP2-GFP ES cells, stained for a neuronal marker NeuN and astrocyte marker GFAP. Scale bar: 50 µm

Neurons derived from ES cells have been used to test the influence of mutations in MeCP2 on its properties. One of the features of wild-type MeCP2 is its localisation in the DAPI-birght foci, consisting of methylated pericentromeric heterochromatin. The neurons produced from 3 different genotypes of ES cells (MeCP2-GFP [WT], MeCP2-GFP [T158M], MeCP2-GFP [R306C]) were fixed and images were collected under the confocal microscope. As expected, the signal from wild-type MeCP2-GFP overlapped with DAPI bright spots (Fig. 5.3.3A). The R306C mutant was also localised to heterochromatin (Fig. 5.3.3B). Conversely, the GFP signal in cells expressing MeCP2-GFP [T158M] mutant was more diffuse and not concentrated in the usual sites of MeCP2 (Fig. 5.3.3C). This observation confirms that a mutation in the methyl-CpG-binding domain of MeCP2 disrupts its localisation on chromatin, and this defect might underlie Rett syndrome caused by the cluster of mutations in MBD.

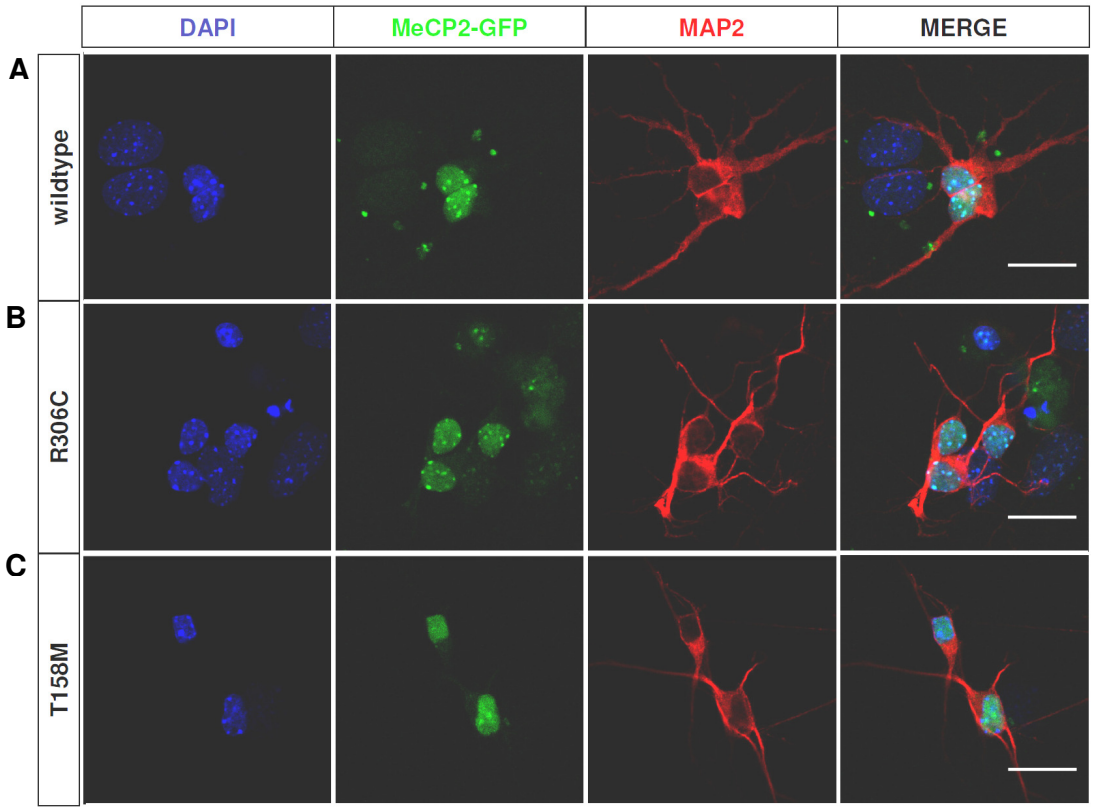


Figure 5.3.3. Neurons derived from ES cells harbouring different mutations in MeCP2 exhibit altered localisation of MeCP2. Confocal microscopy of ES cells differentiated into neurons. DAPI, GFP signal and immunocytochemistry for neuronal marker MAP2 in: (A) MeCP2-GFP [WT], (B) MeCP2-GFP [R306C], (C) MeCP2-GFP [T158M]. Scale bar: 20 µm.

Immunocytochemistry performed using antibodies recognising the neuronal marker, microtubule-associated protein 2 (MAP2), demonstrated that the observed cells were mature neurons (Fig. 5.3.3).

Neurons obtained from ES cells in the same way were used to test the interactions between the endogenous MeCP2 and NCoR/SMRT complex. The effect of mutations in MBD and TRD domains was opposite to their influence on MeCP2 binding to chromatin. GFP immunoprecipitations showed that MeCP2-GFP [T158M] was able to bind the components of the co-repressor complex as well as MeCP2-GFP [WT], whereas the R306C mutation abolished the interaction in neurons (M. Lyst, personal communication). These results corroborate the data obtained in HeLa cells over-expressing MeCP2 mutants, but in a more physiologically relevant system of neuronal cultures with MeCP2 mutations knocked-in under the endogenous MeCP2 promoter.

To summarize, our discoveries made in neurons derived from ES cells allowed us to form a hypothesis that MeCP2 can function as a 'bridge' between methylated chromatin and the NCoR/SMRT complex. Disruption of any of its domains by RTT mutations affects either its proper localisation on chromatin (T158M mutation in MBD) or the interaction with a co-repressor complex (R306C mutation in TRD). The functional consequence of the failure of bringing the co-repressor complex to MeCP2 binding sites in chromatin is not yet known. Therefore, we employed the neuronal differentiation system again, to test the influence of the Rett syndrome-causing mutations on the transcriptional repression and compare the results with *Mecp2* knock-out neurons.

5.4 Transcriptional repression in ES cell-derived neurons

In order to test the hypothesis that MeCP2 functionally links methylated chromatin and NCoR/SMRT co-repressor complex we sought for a system to assess the methylation-dependent transcriptional repression. As there are no convincing genomic targets of MeCP2 identified thus far, we decided to use the methylated reporter assay. It utilises a methylated luciferase gene and relies upon the ability of methyl-CpG binding proteins (MBD) to bind to the methyl groups on DNA. To assess the transcriptional repression function of the tested MBD protein, the luciferase activity recovered from methylated plasmid is then compared with activity recovered from an unmethylated version of the gene. This system has been used in the laboratory previously to assess the methylation-dependent transcriptional repression in cell lines co-transfected with *Mecp2*, *Mbd2* or *Mbd3* genes (Guy et al., 2001; Hendrich et al., 2001). However, it has never been done in the neurons expressing endogenous, mutated versions of MeCP2.

As a first step in the methylated reporter assay the plasmids with the luciferase reporter gene had to be methylated. Preliminary experiments showed very low transfection efficiency of mature neurons and hence low luciferase signals from a plasmid bearing *Firefly* luciferase. Therefore, I decided to methylate the *Renilla* luciferase gene, as it generally gives more luminescent signal and its expression therefore is easier to detect (data not shown). The plasmid coding less active *Firefly* luciferase was used in an unmethylated state as a transfection control (Fig. 5.4.1A).

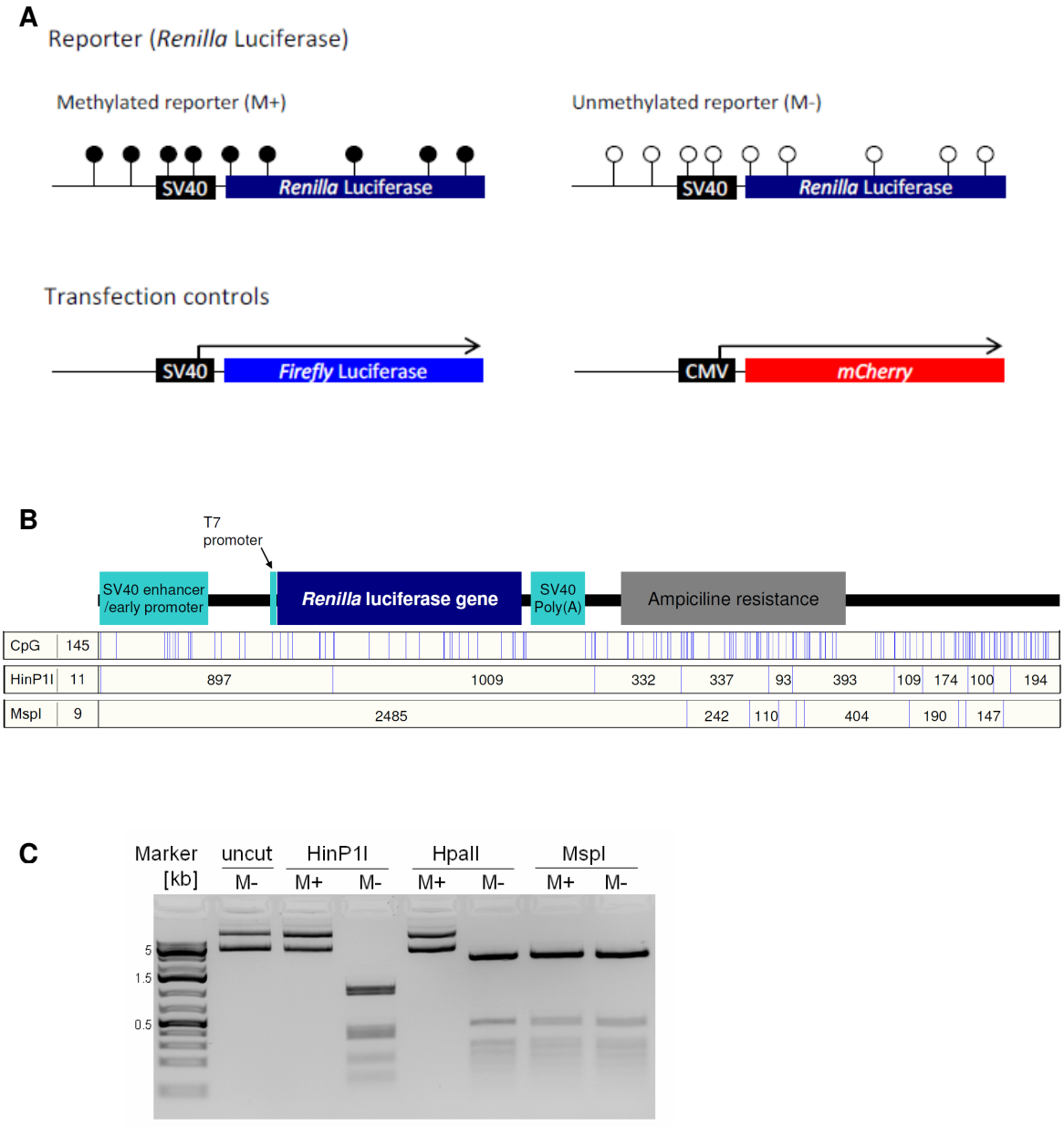


Figure 5.4.1. DNA constructs used in methylated reporter assays. (A) Schematic diagram of the DNA constructs used in methylated reporter assays. (B) The density of CpG dinucleotides on a reporter gene sequence and predicted sizes (in base pairs) of largest fragments given by digestion with HinP1I and MspI/HpaII restriction endonucleases. (C) Confirmation of methylation of SV40-Renilla plasmid by digestion with methylation sensitive restriction enzymes. Agarose gel stained with ethidium bromide. M+ - plasmid methylated *in vitro*, M– - plasmid 'mock' methylated, processed in the same way as M+, but without adding the SssI CpG methyltransferase enzyme.

To assess the effectiveness of *in vitro* methylation with SssI CpG methyltransferase, the processed plasmid was digested by methylation sensitive restriction enzymes and resolved on agarose gel. Both HinP1I and HpaII enzymes failed to cut the methylated plasmid, while they efficiently digested the unmethylated plasmid to the fragments of predicted size (Fig. 5.4.1B and C). To test if methylation and purification processes did not inhibit the restriction enzymatic reaction *per se*, both plasmids were incubated with MspI endonuclease, an isoschizomer of HpaII, insensitive to DNA methylation. In this case both methylated and unmethylated plasmids were efficiently digested. These experiments allowed me to conclude that the effector plasmid was completely methylated at the CpG dinucleotides and therefore could be used in the methylation repression studies.

In order to establish the best method of transfecting mature neurons I performed preliminary experiments testing various gene delivery protocols. A plasmid bearing the mCherry gene was introduced into ES cell-derived neurons using two dilutions of five different transfection agents (JetPEI; X-tremeGENE 9; X-tremeGENE HP; Lipofectamine Plus; and calcium phosphate; see Chapter 2 for details). Live cell imaging revealed that, although very low, the highest transfection efficiency was obtained with the X-tremeGENE 9 reagent (data not shown) and this was therefore used in all subsequent experiments. Staining of fixed cells with DAPI allowed calculation of transfection efficiency under the confocal microscope and it was between 2-3% (Fig. 5.4.2). As the lipid-based gene delivery methods are most efficient against the dividing cells we feared that the transfected cells, although expressing MeCP2-GFP (Fig. 5.4.2), were not neurons. However, co-staining the cells with anti-MAP2 antibody revealed that approximately 20% of transfected cells were neuronal (Fig. 5.4.3). The other transfected cells could be immature neurons, glial cells or undifferentiated ES cells.

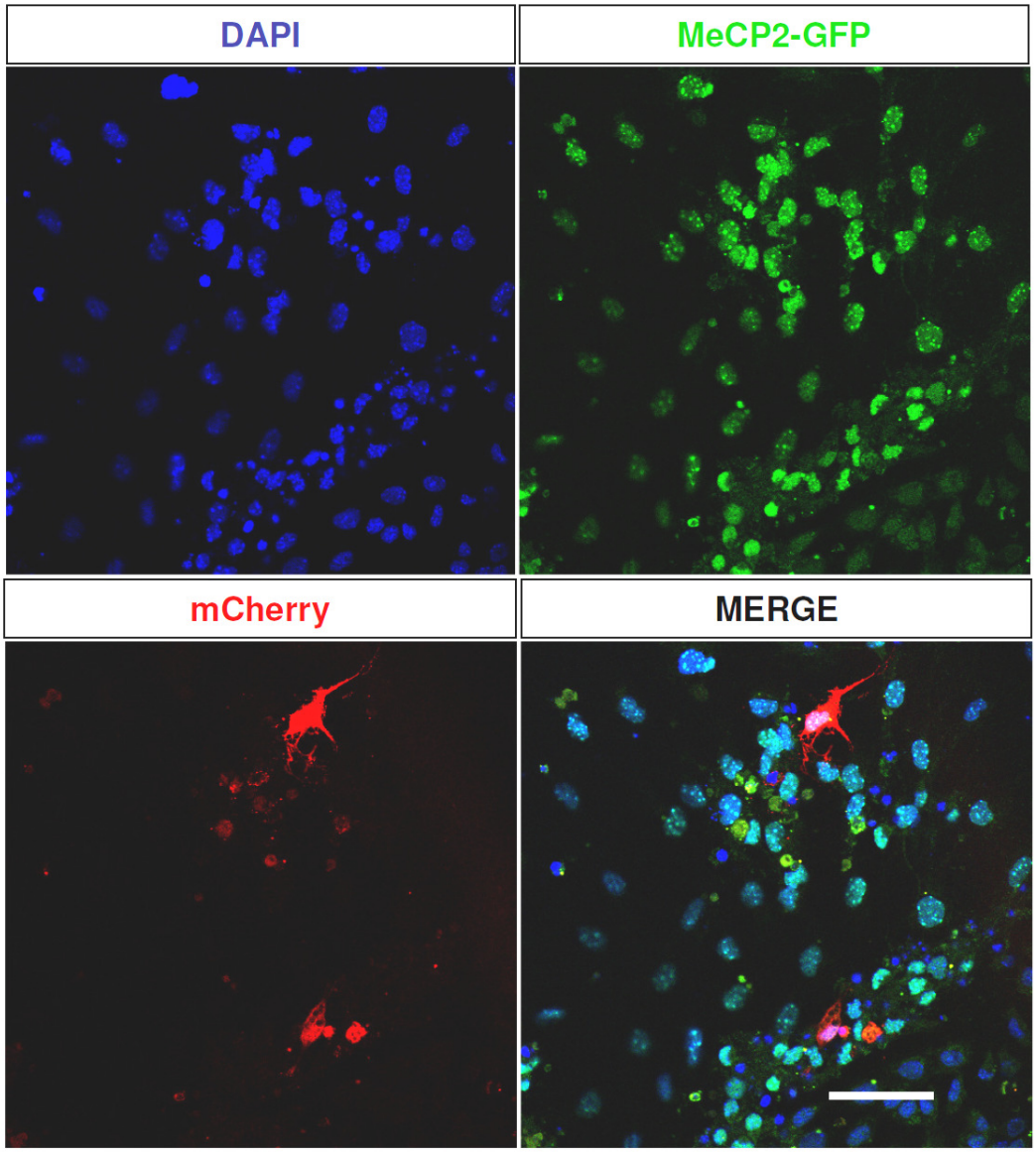


Figure 5.4.2. Transfection efficiency of ES cell-derived neurons. Representative images of ES cell-derived neurons transfected with a methylated reporter gene and co-transfected with a plasmid encoding mCherry protein. Scale bar: 50 µm.

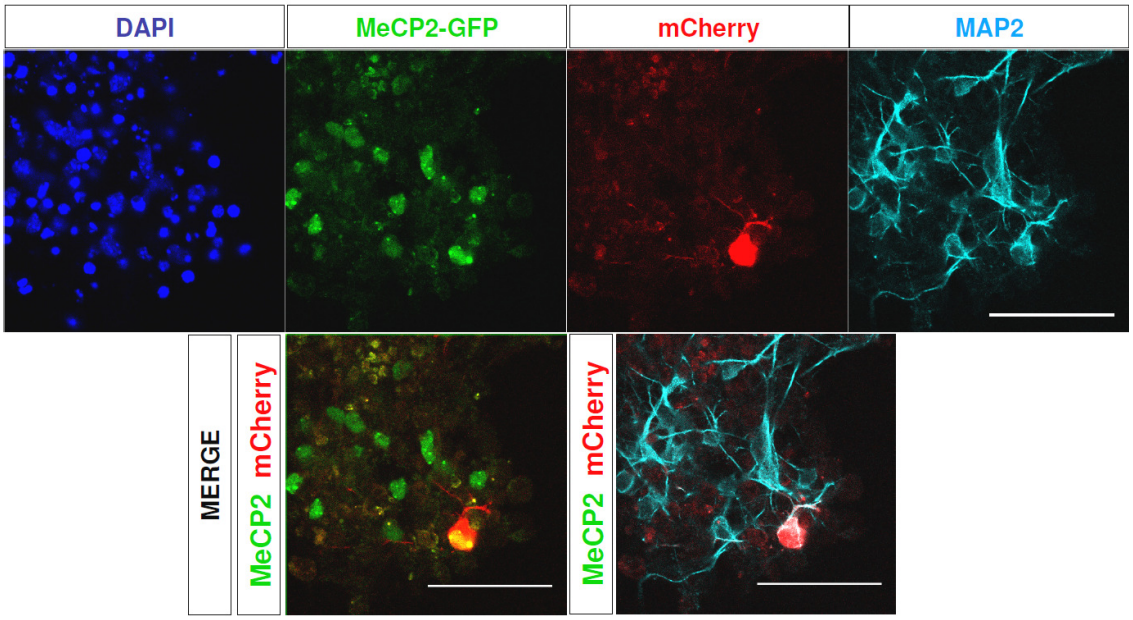


Figure 5.4.3. Transfection of neurons. Imunocytochemistry with anti MAP2 antibody revealed that some of the cells successfully transfected with plasmid coding mCherry protein were neurons. Scale bar: 50 µm.

Five genotypes of neurons, two expressing wild-type MeCP2 (wild-type ES cells and MeCP2-GFP [WT] cell line) and three with alterations in the *Mecp2* gene (knock-out, MeCP2-GFP [T158M] and MeCP2-GFP [R306C]) were then produced. Cells were transfected with luciferase plasmids and the effect of MeCP2 on reporter gene expression was assessed by luminometry. Fold repression was calculated as a ratio of normalised luminescence in cells transfected with unmethylated reporter gene over the normalised luminescence in cells transfected with methylated plasmid. The means were calculated from the values obtained in five independent biological replicate experiments. Overall, there were no significant differences between fold repression of methylated reporter genes in the five tested neuronal genotypes (Fig. 5.4.4). The methylated gene was expressed on average 7.6-fold lower level than its unmethylated counterpart, regardless of the MeCP2 presence or mutation. A weak trend towards higher repression of luciferase gene in cells expressing wild-type MeCP2 than in *Mecp2* knock-out cells could be observed, but it was not statistically significant (unpaired two-sample Student's t-test assuming equal variances; $p > 0.05$).

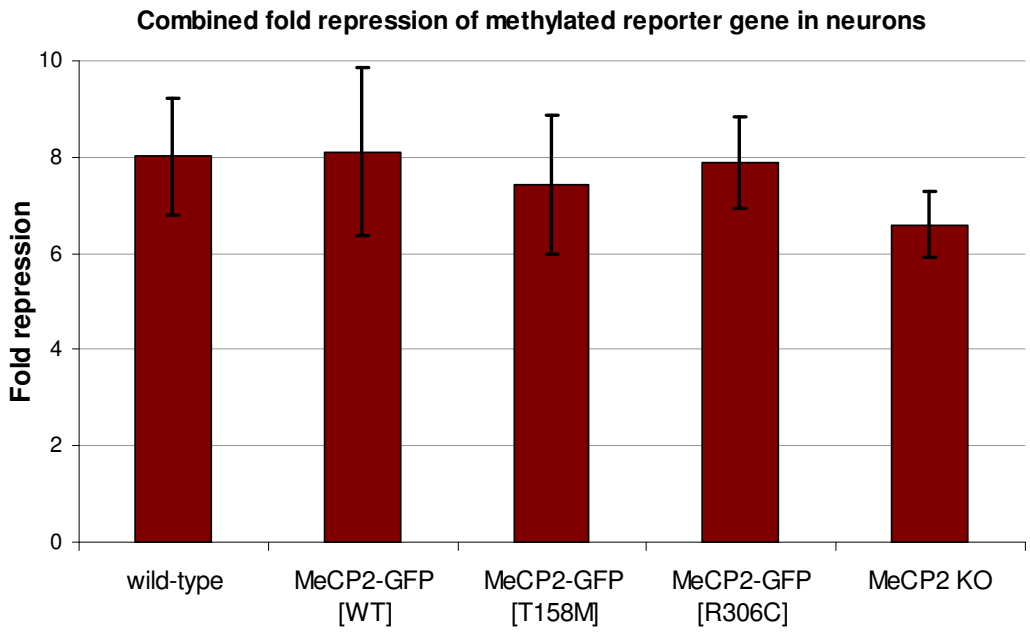


Figure 5.4.4. Mutations in MeCP2 do not influence the repression of methylated reporter gene in ES cell-derived neurons. Combined fold repression of methylated reporter gene over its unmethylated counterpart in neurons derived from ES cells expressing different types of MeCP2 or with *Mecp2* gene knocked out (KO). Average of five biological replicate experiments. Error bars represent standard errors of the means.

5.5 Lower protein levels of mutated MeCP2 in ES cell-derived neurons

As the methylated reporter assay used here depends on endogenous proteins, we hypothesised that the negligible effect of MeCP2 on methylated reporter gene expression may be due to the fact that the neurons used in these experiments were not mature enough and were therefore expressing inadequate levels of MeCP2. Although the confocal microscopy analysis revealed green signal in the neuronal cultures of cells with GFP-tagged MeCP2 knock-in, it was not quantitative and it was difficult to compare the MeCP2 levels between the cell genotypes. To investigate the amounts of MeCP2, protein extracts from ES cell-derived neurons were analysed by Western blot. MeCP2 was detected in all but *Mecp2* knock-out neurons. Two wild-type versions of the protein (endogenous MeCP2 in wild-type ES cells and the

knocked-in MeCP2-GFP fusion protein) were expressed at similar levels (Fig. 5.5A). Interestingly, in neurons harbouring two different missense mutations in *Mecp2* (T158M and R306C) the protein levels were lower than in wild-types by densitometric analysis (Fig. 5.5B). Similar phenomenon of lower levels of endogenously expressed mutated MeCP2 was recently observed in the *Mecp2*$^{T158A/y}$ mouse (Goffin et al., 2012). The potential significance of these discoveries is discussed later.

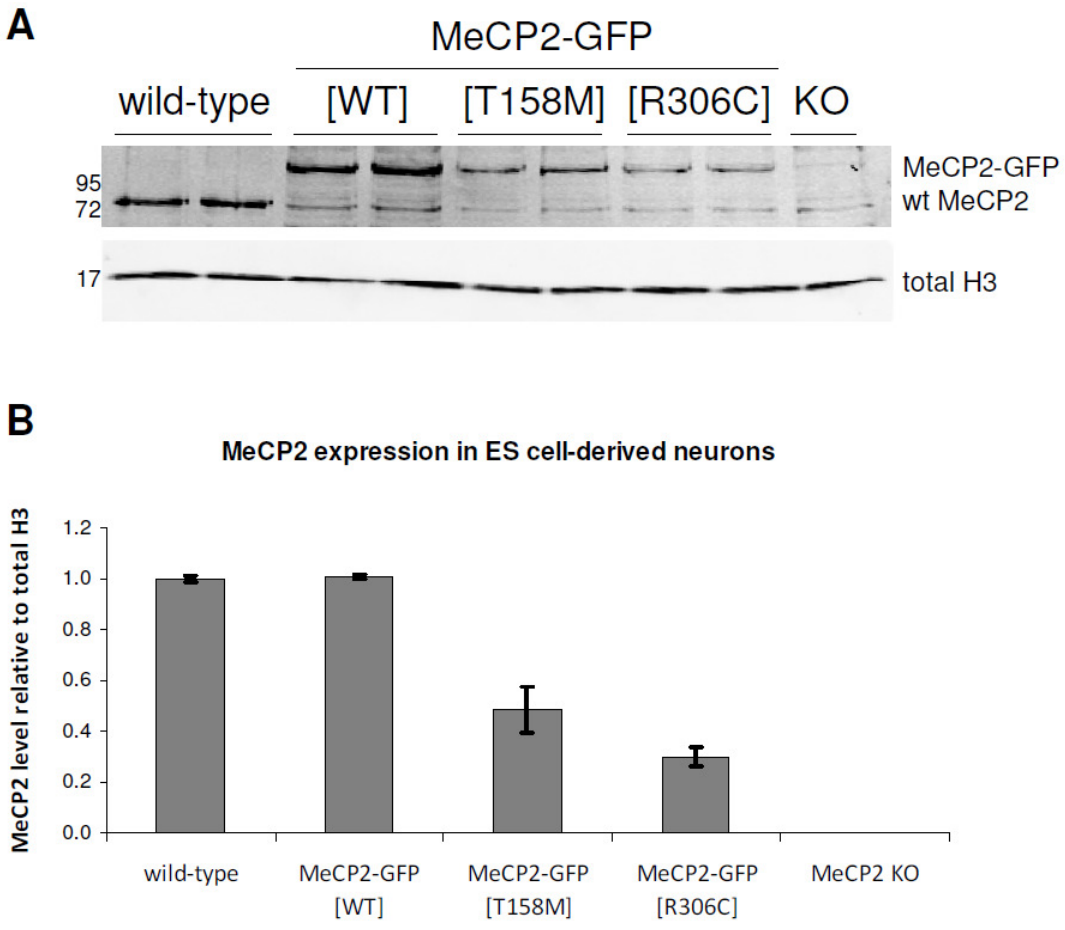


Figure 5.5 Expression of MeCP2 mutants in the ES cell-derived neurons.
(A) Representative Western blot of protein extracts from ES cell-derived neurons probed with anti-MeCP2 antibody and anti-H3 as a loading control. (B) Densitometric analysis of MeCP2 protein levels relative to histone H3 and normalised to wild-type cells.

5.6 Immunohistochemistry of HDAC3 and MeCP2

The functional interaction between MeCP2 and the NCoR/SMRT complex suggests that the complex may be targeted to chromatin at the MeCP2 binding sites. The very high abundance of MeCP2 molecules in neurons, reaching almost the number of nucleosomes, results in coating whole chromosomes while tracking the methylated CpG dinucleotides (Skene et al., 2010). Microscopy images of mouse neurons expressing MeCP2-GFP fusion proteins show that it localises in the DAPI bright spots consisting of highly methylated pericentromeric heterochromatin. We hypothesised that the components of NCoR/SMRT complex should also localise in the spots occupied by MeCP2, and that this localisation should be disrupted upon MeCP2 deletion. In order to test this hypothesis I used a brain from a heterozygous *Mecp2-GFP*$^{+/-}$ female mouse for immunohistochemistry, as described in Chapter 4.4.1. As expected, only half of the neurons were MeCP2 positive and the GFP signal from MeCP2 fusion was clearly co-localised with the DAPI bright spots (Fig. 5.6A and B). The immunofluorescent staining with anti-HDAC3 antibody gave a strong signal in the nuclei, but it seemed to be excluded, rather than enriched, from the heterochromatic foci. The HDAC3 distribution appeared not to be different between MeCP2 positive and knock-out neurons (Fig. 5.6A, insets). This observation was true in four pictures from two different brain slices, with ~50 neuronal nuclei analyzed on each. The obtained fluorescent signal was not due to unspecific binding of the secondary antibody as the control staining omitting primary antibody gave no signal at all (Fig. 5.6B). Immunofluorescent staining with antibodies against other components of NCoR/SMRT complex was either very weak or gave high background and was therefore impossible to interpret (data not shown). These are only preliminary results and they do not disprove the formulated hypothesis until they are systematically repeated with the antibodies with checked specificity or with other optimised protocols for immunohistochemistry.

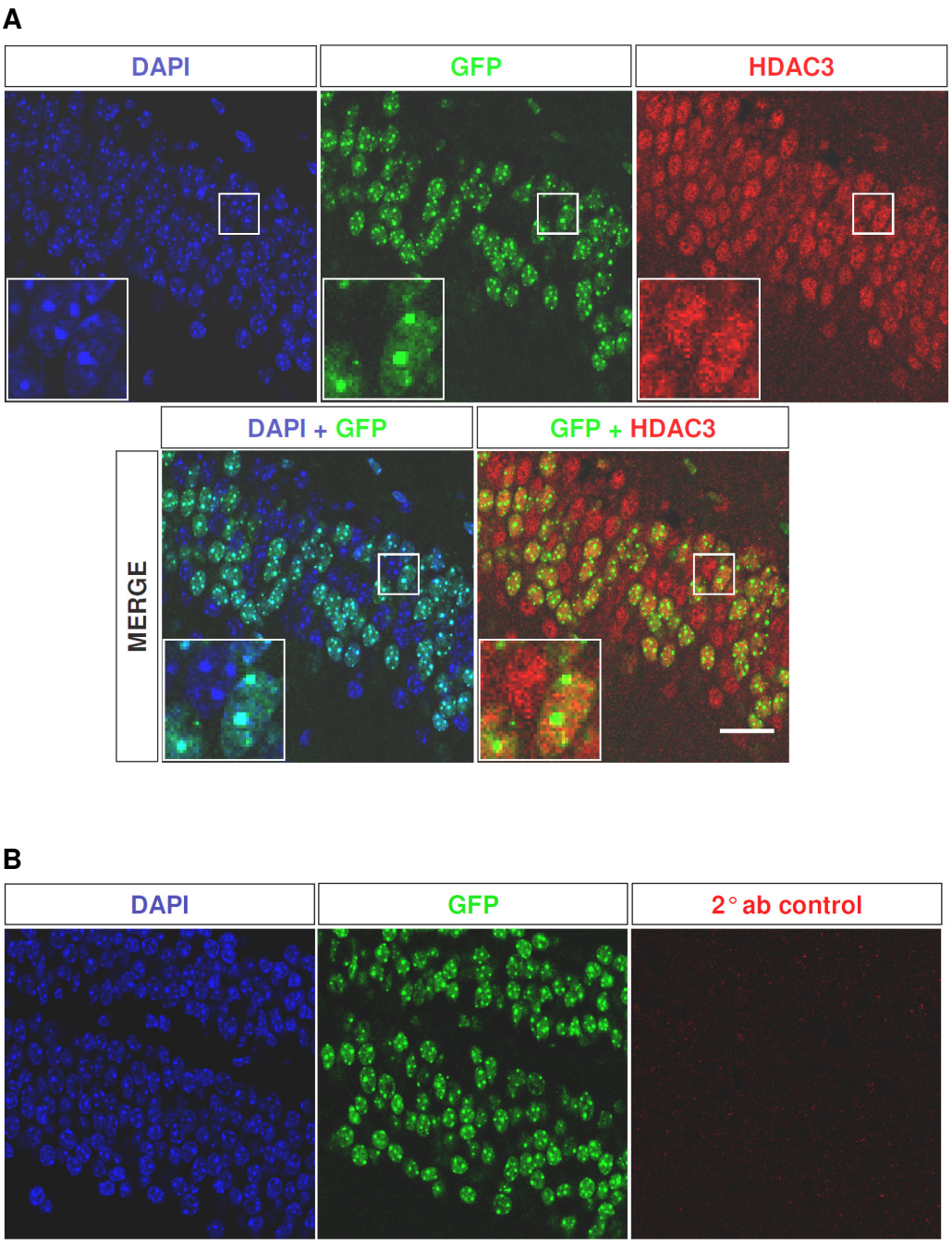


Figure 5.6. HDAC3 does not co-localise with MeCP2 in mouse neurons. Confocal microscopy of *Mecp2-GFP*$^{+/-}$ heterozygous female mouse brain. (A) Immunohistochemistry for HDAC3. Insets are a 3x magnification of a region marked with white square. Representative images of 2 brain slices (B) Staining control with secondary antibodies only, omitting the primary antibody step. Scale bar: 25 µm.

5.7 Discussion

In this chapter I showed that the identified interaction between MeCP2 and NCoR/SMRT co-repressor complex is probably crucial for MeCP2 function as a transcriptional repressor and the disruption of this function likely plays a role in Rett syndrome. The NCoR/SMRT complex binds to a fragment of MeCP2 containing residues commonly mutated in RTT, including the second most frequent missense mutation, R306C (RettBASE). The identified NCoR/SMRT-interacting region on MeCP2 is quite wide (40 aminoacids between positions 269-309) and needs to be defined more precisely in the future. The fact that A279V mutant binds the complex and is able to repress transcription suggests that the minimal region of interaction may be shorter. Also, the biotinylated peptide comprising of 285-319 amino acids was sufficient to interact with the co-repressor complex in the *in vitro* binding studies (J. Nowak, personal communication).

The single amino acid substitutions in the identified region that cause Rett syndrome prevent binding of the complex and also lead to the impairment of transcriptional repression by MeCP2. The results presented in this chapter are the first example of a functional consequence of a disease-causing missense mutation in the TRD of MeCP2. Two previous reports claimed that MeCP2 [R306C] mutant can repress transcription as efficiently as the wild-type protein (Yusufzai and Wolffe, 2000; Drewell et al., 2002). Our contrasting results are most probably due to the use of saturating amounts of MeCP2 by other laboratories. It also suggests that this mutant possesses some residual repression activity, possibly due to residual binding of NCoR/SMRT complex. Although we could not detect it by co-immunoprecipitations from both HeLa cells and ES cell-derived neurons expressing MeCP2 [R306C], there is one report demonstrating weak binding of SMRT to this mutant (Stancheva et al., 2003). Another possibility is that the residual repression activity of MeCP2 [R306C] is mediated by Sin3A complex, which binds to a different region of TRD (Nan et al., 1998) and the RTT mutations in the 302-306 cluster do not affect it. The interaction of MeCP2 with Sin3A, though, is more than 10-fold weaker than the interaction with NCoR/SMRT (M. Lyst, personal communication). It suggests that although real, the interaction of MeCP2 with Sin3A

is not the major one and its disruption is not responsible for RTT. The evidence that RTT mutations in 302-306 cluster abolish binding of NCoR/SMRT co-repressors and cause significant impairment of the repression properties (at least in cell culture systems) is much stronger and should change the way we look at MeCP2.

We of course bear in mind that our discovery was made in a very artificial system, where MeCP2 is fused to the GAL4 DNA binding domain, which targets it directly to the reporter gene. To fully appreciate the importance of our findings it would be helpful to be able to examine a mouse bearing the R306C mutation in MeCP2. The prediction is that such mouse should recapitulate the entire spectrum of Rett syndrome phenotypes making it invaluable tool to study the pathology of the disease. To ultimately confirm that the interaction between MeCP2 and NCoR/SMRT lies in the basis of the disease, a system that allows rescuing the abolished interaction is also needed. For example suppressor mutation on NCoR/SMRT or GFP-binding domain fused to NCoR to enforce the interaction with MeCP2-GFP [R306C]. If such a mouse has the Rett phenotype rescued, then it may be worth screening for drugs that permit the co-repressor binding to the mutated MeCP2.

Regrettably the R306C mouse is not available yet, therefore we tried to dissect the molecular function of MeCP2 by means of other methods. The embryonic stem cell differentiation system appealed as a good way of obtaining pure neuronal cultures with various mutations in MeCP2 due to the ease of introducing mutations in ES cells. Indeed, the procedure of neuron production was successful and showed that wild-type MeCP2-GFP expressed from endogenous locus localised correctly to heterochromatin. The localisation of MeCP2-GFP [R306C] was indistinguishable from the wild-type protein, but the MeCP2-GFP [T158M] distribution on chromatin was more diffuse. This mutation was previously shown to disrupt the MeCP2 structure and in consequence its interaction with methylated DNA in the *in vitro* binding assays (Yusufzai and Wolffe, 2000; Ghosh et al., 2008; Ho et al., 2008). Also the over-expression of the MeCP2 [T158M] in cell cultures proved its faulty localisation to heterochromatin when compared to wild-type protein (Kudo et al., 2003; Schmiedeberg et al., 2009). This is the first time when the MeCP2 [T158M] mutant expressed from endogenous locus in neurons is shown to have a binding

defect. It supports the data from Zhou laboratory, where MeCP2 [T158A] mouse also exhibited mislocalisation of MeCP2 (Goffin et al., 2012).

All of the mentioned localisation studies were performed using fixed cells or brain sections. However, when T158M and other MeCP2 mutants in MBD were over-expressed in NIH-3T3 cells, live cells imaging showed their normal localisation (Schmiedeberg et al., 2009). Also the neurons produced from MeCP2-GFP [T158M] ES cells showed MeCP2 enriched in heterochromatic foci to some extent when observed in live cells (data not shown). Nevertheless, as the cell fixation procedure does not disrupt the localisation of wild-type MeCP2, it is still safe to say that MeCP2 [T158M] has some binding defect. It may be due to the altered binding dynamics (Kumar et al., 2008) and should be investigated in further details once the MeCP2-GFP [T158M] mouse is available.

The subsequent experiments that followed the ES cells differentiation into neurons failed to find the functional consequences of RTT mutations in *Mecp2*. Luciferase reporter assays performed in neurons harbouring different mutations of MeCP2 did not show any significant differences in transcriptional repression of the methylated reporter gene between wild-type MeCP2, mutants and knock-out. One of the possible interpretations is that the primary function of endogenous MeCP2 is not the transcriptional repression, which was observed previously mostly in the cultured cells transfected with engineered MeCP2 constructs. However there are also other, technical, explanations. The luciferase signal in cells transfected with methylated plasmid was always lower than in case of a plasmid free of methylation (~7.6 fold). This may be due to the presence of other methyl-CpG binding proteins in neurons, higher compaction of the chromatinised methylated plasmid, or the simple occlusion of transcription factors by the methyl groups on CpG dinucleotides. All of these factors could mask the effects of mutations in MeCP2 on the reporter gene expression and contribute to the fact that there were no significant differences identified. It is also possible that the amounts of plasmid coding the reporter luciferase and its ratio to the control plasmids were not chosen optimally. If the number of molecules of reporter gene DNA was in a large excess over the amount of available MeCP2 molecules, then the effect of mutations in the protein would be undetectable as the majority of luciferase signal would come from the excess of its

gene unbound by MeCP2. Unfortunately, due to the nature of these experiments and very time-consuming process of neuron production, the testing of many various plasmid ratios and a systematic optimization of the assay was impossible in the available time.

Although MeCP2 was detected in the ES cell-derived neurons it is possible that the Western blot analysis revealed average levels of MeCP2 in the population of cells. If the subset of cells that were transfected contained smaller amounts of MeCP2, it could partially explain why the methylated reporter assay in these cells failed to show any differences between the MeCP2 genotypes. Methylated plasmid has to compete for MeCP2 binding with the whole genome. The components of NCoR/SMRT co-repressor complex could also not be expressed at the same levels in the transfected ES cell-derived neurons as in the mature brain and therefore, being a limiting resource, masking the effects of mutations in MeCP2.

If these experiments were to be continued the priority would be to ensure that the neurons used are mature and express uniformly large amounts of the tested proteins. This could be achieved by keeping them in culture for longer periods of time, however the preliminary studies showed that there is an increased risk of cell death during prolonged *in vitro* tissue culture.

Improvement of the transfection efficiency is also essential, as it would allow scaling down the whole procedure and in consequence enable the proper optimization of used DNA ratios. The usual method of gene delivery into non-dividing cells is viral transduction, but unfortunately it cannot be used in this case where the plasmid is methylated since the methylation marks would be removed during the viral life cycle. Instead of inefficient chemical-based methods of neuron transfection, maybe cells could be electroporated. The standard procedure that involves cell trypsinisation was utilized to transfect luciferase constructs into primordial neural stem cells and achieved ~30% transfection efficiency, which was enough to detect luciferase activity in used assays (Muotri et al., 2010). However, this approach cannot be used in the ES cells differentiation system for testing the effects of endogenous MeCP2, because transfection at the progenitor stage would be too early to reach neuronal maturity and trypsinisation of cultured neurons at later stages would destroy the neuronal network and connections. However, there are *in*

situ electroporators available that can deliver DNA to the adherent cells. Purchase and use of such equipment should be considered in the future.

The lower levels of MeCP2-GFP [T158M] and MeCP2-GFP [R306C] mutants than the wild-type versions of the protein could be an effect of impaired neuronal maturation when MeCP2 is mutated. Against this hypothesis is the fact that neurons of all genotypes express neuronal markers NeuN and MAP2. Another possibility could be faster degradation of mutated proteins. This would stand in line with recent discovery that in a mouse model harbouring a MeCP2 [T158A] mutation the MeCP2 protein stability is decreased (Goffin et al., 2012). In the same publication fibroblast cultures derived from RTT patient with T158M mutation also showed reduced level of MeCP2 when compared to an age-matched female control. If this phenomenon is also true for R306C, and perhaps other mutations, it could make a contribution to the molecular basis of Rett syndrome. Mutated proteins can be misfolded, recognized by the protein degradation machinery and destined for proteolysis. This hypothesis would explain why missense mutations outside of the functional domains of MeCP2, as well as nonsense mutations in the C-terminus can cause Rett syndrome. This mechanism can be especially true in case of P255R and P322L mutations, where proline residue has been substituted with a structurally distinct amino acid. As proline has a very rigid conformation (Fasman, 1989) it is possible to imagine how its substitution can affect proper protein folding. At present, however, there is no evidence relating to these hypotheses.

To our surprise immunohistochemistry of one of the components of NCoR/SMRT complex, HDAC3, did not show the co-localisation with MeCP2 in *MeCP2-GFP*$^{+/-}$ brain slices. Differences in HDAC3 distribution in neurons expressing and lacking MeCP2 have also not been detected. These results may be explained by technical constrains of the immunohistochemistry procedure itself. MeCP2 localises in the heterochromatin and such environment might not be permissive for the anti-HDAC3 antibody. Use of other antibodies or one of the antigen retrieval methods could improve the staining and give different results. It is also possible that the interaction between MeCP2 and NCoR/SMRT complexes is very transient and dynamic. The fast turnover would be enough to repress

transcription wherever MeCP2 is bound, but it is not sufficient to detect the co-localisation in a snap-shot of fixed cells. As the number of MeCP2 molecules in neurons is reaching the number of nucleosomes (Skene et al., 2010) the abundance of components of NCoR/SMRT complex is probably much lower. The complex has other functions independent of MeCP2 and its own binding sites could be occupied by it and remain unchanged regardless of MeCP2 status. Therefore, the lack of expected results in these experiments does not rule out the importance of the interaction. The same limitations may apply to other techniques used or planned to be used to investigate the interaction further, such as chromatin immunoprecipitation. ChIP of HDAC3 or other components of NCoR/SMRT complexes followed by high-throughput sequencing could tell us how many MeCP2 binding sites are also occupied by the repressor complex. Unfortunately, preliminary results suggest that cell fixation with paraformaldehyde disrupts the MeCP2 interaction with NCoR/SMRT complex, as assessed by immunoprecipitations followed by Western blotting (M. Lyst, personal communication). The 269-309 region of MeCP2 identified as a binding surface for NCoR/SMRT complex contains 8 lysine residues in close proximity (Fig. 5.7). These lysines may be cross-linked to each other instead of fixing the interaction between MeCP2 and proteins interacting in that region. This could also be the main reason why HDAC3 immunohistochemistry of fixed cells showed no co-localisation with MeCP2 and even the exclusion from MeCP2-dense sites. Use of a different cross-linking agent may have to be necessary in future experiments.

Figure 5.7. Lysine residues in the C-terminal end of TRD domain. 269-309 fragment of mouse MeCP2 isoform e2. Known RTT mutations are underlined, arrows indicate lysine residues.

Chapter 6: Final conclusions and future perspectives

This study provides some new insights into the complex matter of the function of methyl-CpG binding protein MeCP2. More importantly, it extends the discoveries made in the laboratory regarding the role of MeCP2 in Rett syndrome. Our increased understanding of the molecular basis of this disorder gives hope for developing a successful therapy in the future.

6.1 Novel interaction partners of MeCP2

Numerous studies using various approaches and techniques reported a plethora of proteins possibly interacting with MeCP2 and mediating one of its many putative functions. Due to inconsistencies between many of these reports, we decided to revisit the issue by taking advantage of recently developed molecular tools. Our physiological and unbiased search for interacting partners of MeCP2 in mouse brains revealed a list of candidate proteins of potential interest.

Amongst the proteins co-immunoprecipitated with MeCP2 we detected five out of six known core components the NCoR/SMRT co-repressor complex. Based on candidate-approach studies two of these proteins have been reported previously to interact with MeCP2 (NCoR and SMRT), but these discoveries were not followed up (Kokura et al., 2001; Stancheva et al., 2003). We additionally identified and confirmed novel interactions of MeCP2 with other proteins building the complex: TBL1, TBLR1 and HDAC3. This suggested strongly that the whole complex is bound by MeCP2. The functional consequence and significance of these interactions to the Rett syndrome were found in this study and are discussed later.

The other identified and validated interactions of MeCP2, with CK2α and CaMKIIα protein kinases, may explain the mechanism of post-translational modifications of MeCP2. Combining this data with the studies on binding of the

co-repressor complex may aid the future discoveries of how the function of MeCP2 is dynamically regulated.

Two of the novel interactions detected in our search were with proteins belonging to the importin α family. The unusually strong binding of MeCP2 to both KPNA3 and KPNA4 suggested the significance of these interactions. However, I could not find any consequence of these interactions other than their role in the transport of MeCP2 to the nucleus.

There are also other proteins on our list that, if confirmed to be real binding partners of MeCP2, may help in explaining its molecular mechanism of action. However, in light of the gathered evidence for the importance of interaction between MeCP2 and NCoR/SMRT complex, the elucidation of the role of the other proteins does not seem like a priority at the moment.

6.2 Only one NLS of MeCP2 is bound by importins α

After more than 20 year of research on MeCP2 it appears as a very complex protein with multiple interactions and roles in cells. The reports often contradict each other regarding the functional domains of the protein. This study should help at least in resolving the issue of number and position of functional nuclear localisation signals in MeCP2. The existence of at least 2 NLS sequences has been postulated from bioinformatic and localisation studies of MeCP2. However, our data clearly demonstrated that only one NLS (covering the region between 253-269 aa) is bound by importins α, suggesting that the other sequences are dispensable for the active transport of MeCP2 to the nucleus. More experiments are needed to explain the phenomenon of partial nuclear localisation of the truncation mutants of MeCP2 that are not able to interact with importins α. The hypothesis that MeCP2 is being retained in the nucleus due to its high affinity to methylated DNA will be tested in the near future.

Based on the evidence presented in Chapter 4 the function of MeCP2-importins α interactions is probably only in the transport to the nucleus. The lack of MeCP2 did not affect importins localisation, expression or stability. The MeCP2-depleted cells exhibit significantly smaller nuclei, but this effect cannot be

attributed to the strong interaction between MeCP2 and KPNA3 and KPNA4.
A possible link between lack of MeCP2 and nuclear size comes from the demonstration that neurons of *Mecp2*-null mice exhibit 2-fold increase of histone H1 levels (Skene et al., 2010) and H1 was shown to compact chromatin and reduce the size of nuclei (Shen et al., 1995). Although the higher order structure of chromatin was more condensed in the presence of MeCP2 than histone H1 (Georgel et al., 2003; Nikitina et al., 2007a) these studies were performed only *in vitro*, and their relevance *in vivo* is not known. It is also possible, that the smaller nuclei phenotype is not due to direct action of MeCP2, but is a secondary effect reflecting lack of the MeCP2 function elsewhere. A Rett nonsense mutant MeCP2 [294X] would be a good model to study the consequences of MeCP2 interaction with importins. Such a mutant does not interact with NCoR/SMRT co-repressor complex but still binds KPNA3 and KPNA4. If the nuclear size of cells expressing the endogenous MeCP2 [294X] is smaller than of wild-type cells, it would demonstrate that the MeCP2 interaction with importins is not ruling the size of nuclei, but it is an effect of disrupted transcriptional repression mediated by MeCP2-NCoR/SMRT interaction. ES cells harbouring this mutation in endogenous *Mecp2* locus are being generated in the laboratory at the moment.

6.3 NCoR/SMRT is the co-repressor complex responsible for transcriptional repression function of MeCP2, which is abolished by RTT mutations

The identification of the MeCP2 interaction with the NCoR/SMRT co-repressor complex and its subsequent mapping to the C-terminal region of transcriptional repression domain of MeCP2 indicated its possible function in repressing transcription. Moreover, the NCoR/SMRT-binding region on MeCP2 is coincident with a cluster of frequent missense mutations causing Rett syndrome. Further analysis showed that RTT mutations in this region indeed abolished the interaction. In this study I investigated the effect of the various mutations disrupting the MeCP2-NCoR/SMRT complex, on the transcriptional repression of reporter genes. I found that these MeCP2 mutants are not able to efficiently repress transcription. The repression by versions of MeCP2 that could interact with

NCoR/SMRT was dependent on histone deacetylase activity. Future studies should focus on testing a wider spectrum of mutations to better define the region responsible for NCoR/SMRT-mediated transcriptional repression.

The implications from these studies are important for understanding the pathology of Rett syndrome. It is the first time that the interaction of MeCP2 with another protein was shown to be abolished by the RTT-causing mutations. The disruption of this interaction affected the ability of engineered constructs of MeCP2 to repress the transcription of reporter genes. The simplest conclusion is that the NCoR/SMRT-mediated transcriptional repression may be the major mechanism of function of MeCP2.

Although very probable from the gathered evidence, the link between other RTT mutations in the full-length protein and methylation-dependent transcriptional repression has not been shown directly *in vivo*. The ability of a wild-type MeCP2 to repress the methylated reporter gene was demonstrated previously (Guy et al., 2001) and surprisingly the R306C mutant was shown to be fully functional (Drewell et al., 2002). This effect could be attributed to the over-expression of MeCP2 masking its impairment in repressing transcription, as we demonstrated that R306C exhibits some residual repression activity. Therefore, we tried to address this issue by testing the effect of RTT mutations in the endogenous *Mecp2* gene expressed in the ES cell-derived neurons. Unfortunately, due to technical problems the experiments were inconclusive. Perhaps expressing MeCP2 constructs in the transfected cell line, although less physiologically relevant, would be an easier approach. The amounts of plasmids used will have to be carefully titrated in order to avoid the problem of masking the detrimental effect of mutations by saturation of the system with over-expressed proteins. In principle such a system would allow testing of the effects on methylation-dependent transcriptional repression of mutations throughout the whole length of MeCP2.

6.4 Re-defining the functional domains of MeCP2

In this study I ruled out the existence of functional active NLS sequences other than the 253-269. Our recently obtained data regarding MeCP2 binding to

NCoR/SMRT complex and the necessity of this interaction to mediate transcriptional repression suggest that we should also change the way we look at the TRD of MeCP2. The transcriptional repression domain identified 15 years ago (Nan et al., 1997) is probably not a single functional entity. The binding to NCoR/SMRT occurs only at the C-terminal end of the domain (residues 269-309). More detailed mapping of the interaction in the future should allow identification of even smaller minimal binding region. Why is the TRD domain so long then, and why are the amino acids 207-269 necessary for proper transcriptional repression (Nan et al., 1997)? If the N-terminal region of TRD interacts with Sin3A it is possible that these two complexes collaborate with each other for the better repression effect. Supporting this idea is the fact that NCoR was reported to be able to interact with Sin3A *in vitro* and *in vivo* (Alland et al., 1997; Heinzel et al., 1997; Wang et al., 1998; Jones et al., 2001). Against this hypothesis stands our data that the interaction of MeCP2 with NCoR/SMRT subunits is much stronger than with Sin3A complex. Additionally, the cluster of frequent mutations causing Rett syndrome disrupts the interaction with NCoR/SMRT, but not with Sin3A, and the mutations with opposite effect have not been identified. At the moment we do not have an explanation why the TRD of MeCP2 is much longer than the surface sufficient to bind to NCoR/SMRT. To address it, we would like to revisit the TRD-mapping studies, using more truncation mutants or the alanine substitutions scanning to re-define the minimal region sufficient for transcriptional repression. Also, more detailed mapping analysis of the interaction on both MeCP2 and NCoR subunits, followed by an attempt to crystalise fragments of the complex to obtain its structure will hopefully help solving this puzzle. A detailed description of the character of MeCP2-NCoR/SMRT interaction, supported by the crystal structure would be a great aid in the design or screening for new drugs and potential therapies for Rett syndrome.

6.5 The levels of mutated MeCP2 in ES cell-derived neurons are lower than wild-type protein

Our efforts in obtaining new tools to study the effects of Rett syndrome mutations on MeCP2 function were successful only to some extent. Generation of

neurons derived from a series of ES cells harbouring various versions of *Mecp2* gene allowed us to investigate the properties of endogenous MeCP2 in a more physiologically relevant environment than in transfected cell-line cultures. We confirmed that in this system the R306C RTT mutation in the NCoR/SMRT-binding region abolishes the interaction, but has no effect on MeCP2 localisation to the methylated foci on chromatin. Conversely, the T158M mutation in the MBD does not affect the binding of co-repressor complex but changes the ability of MeCP2 to bind to chromatin. Unfortunately, other experiments performed in the ES cell-derived neurons did not provide conclusive results, most probably due to the technical issues. Interestingly, I have observed that the protein levels of the mutated versions of MeCP2, T158M and R306C, are lower than the wild-type. It was also reported that another RTT mutation, T158A, has decreased stability in mouse (Goffin et al., 2012). If this effect is true for the other Rett syndrome mutations, it may be an additional contribution to the disease pathology and a potential target for future therapies.

6.6 A model of MeCP2 function and the effect of RTT mutations

Based on the results presented in this study we propose a new model of MeCP2 function relevant to Rett syndrome. In the nucleus, MeCP2 can interact with methylated CpGs on DNA and NCoR/SMRT complex, serving as a 'bridge' between these two. The co-repressor complex targeted to methylated chromatin by MeCP2 deacetylates histone tails, probably through its component HDAC3, and in consequence inhibits transcription (Fig. 6.6A). When MeCP2 is mutated in the MBD, such as in the case of T158M or other frequent mutations in the same cluster, it can no longer properly bind to the methylated DNA. Despite its intact interaction with MeCP2, the NCoR/SMRT co-repressor complex is not correctly delivered to its sites in the genome and transcription is not down-regulated. This is reflected by increased acetylation on histone tails (Fig. 6.6B). A mutation in the second RTT-related cluster of genetic alterations, such as R306C, affects the ability of MeCP2 to bind NCoR/SMRT. Proper localisation of MeCP2 to the methylated DNA is not sufficient

to prevent Rett syndrome, because the co-repressor is not correctly targeted to where it should perform its role (Fig. 6.6C).

Our model is not radically different from what was already considered as one of the functions of MeCP2. However, we think that the key player in the repression of transcription is not Sin3A but NCoR/SMRT co-repressor complex. The fact that interaction with it is disrupted by a cluster of mutations causing Rett syndrome and that the other cluster prevents MeCP2 from binding to methylated chromatin allows us to speculate that this function of MeCP2 is the only one underlying the pathology of RTT. The much less frequent RTT mutations outside of these two clusters may affect protein stability or folding, and this would again result in the lack of bridging methylated DNA and NCoR/SMRT. This hypothesis remains to be formally tested. The other reported interactions of MeCP2, and based on them various putative functions, may be true in *in vitro* assays and in some cases even *in vivo*, but the important issue is how relevant to the Rett syndrome they are. Prior to this study this question was never properly addressed.

The NCoR/SMRT co-repressor complex contains the histone deacetylase HDAC3. Therefore, the functional consequence of MeCP2 'bridging' methylated chromatin and the complex should also manifest in a difference in histone acetylation when MeCP2 is mutated or missing. The global hyperacetylation of histones in MeCP2 knock-out brains was observed only when neurons were purified away from glial cells (Skene et al., 2010). The effect of mutations in MeCP2 on global and local acetylation of histones should be tested in the future when MeCP2 [T158M] and MeCP2 [R306C] mice are available. Fluorescence-activated cell sorting of neuronal nuclei followed by Western blot analysis with anti-acetyl H3 antibodies, as well as ChIP assays for acetylated H3 and MeCP2 should reveal if disrupting just one side of the MeCP2 'bridge' gives the same effect as MeCP2 knock-out.

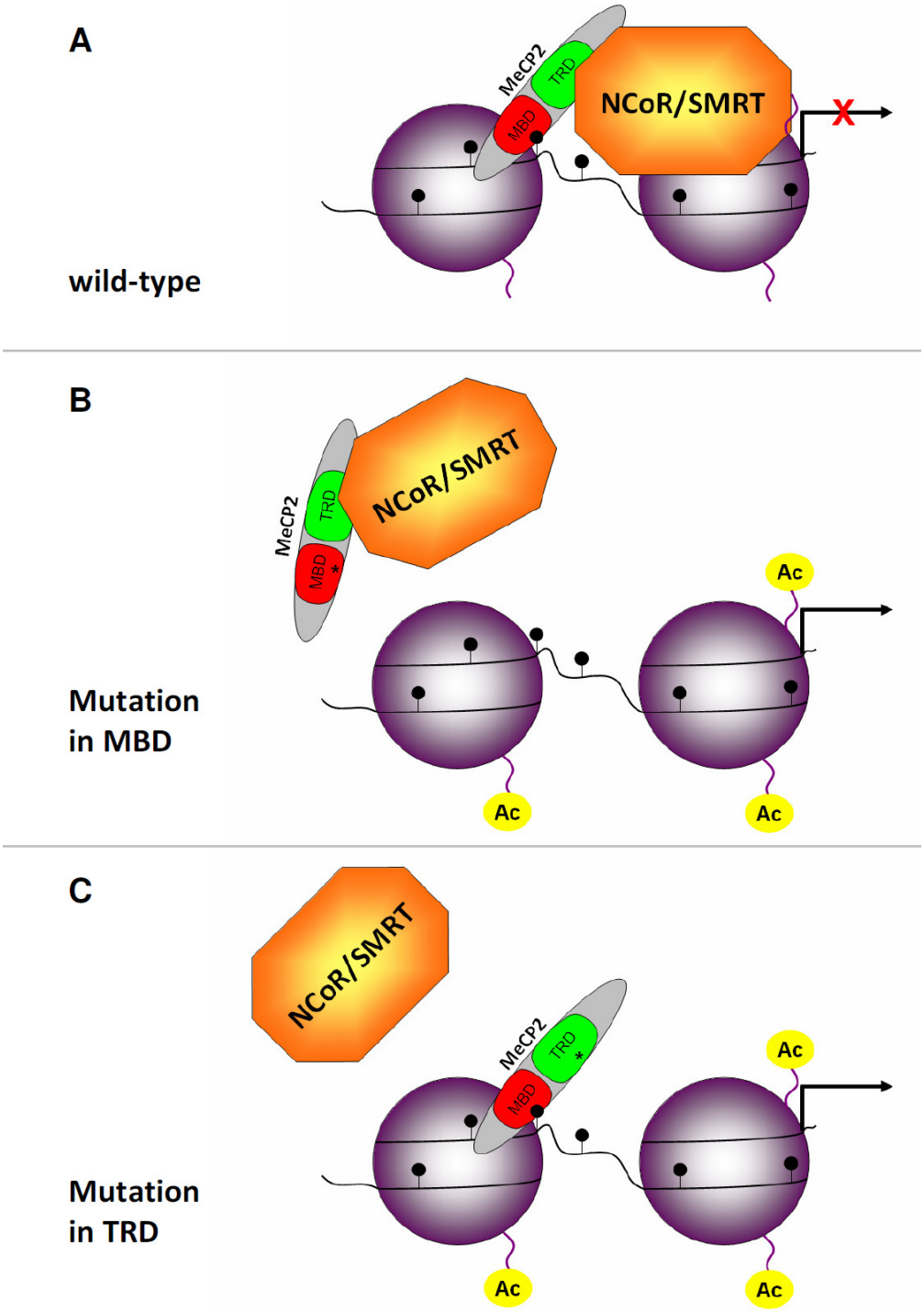


Figure 6.6. A model of MeCP2 function based on the research in this study. The black line depicts DNA wrapped on purple nucleosomes. Black lollipops represent methylated CpGs. Yellow circles labelled 'Ac' are acetylated histone tails. An asterisk in MBD or TRD represents a mutation in MeCP2. (A) In a wild-type cells MeCP2 binds to methylated CpGs and represses transcription by recruiting NCoR/SMRT complexes deacetylating histone tails. (B) When mutated in MBD, MeCP2-NCoR/SMRT can no longer bind to chromatin and repress transcription, resulting in acetylated histone tails. (C) MeCP2 mutated in TRD can bind to methyl-CpG, but fails to recruit NCoR/SMRT and the transcription is not repressed.

Importantly, the targets of MeCP2-NCoR/SMRT-mediated repression are not known. As the genes regulated by MeCP2 have not been convincingly identified, it is possible that MeCP2 affects global transcription, maybe from non-coding regions. The genome-wide search for the consequences of lack of MeCP2 on the transcription is on-going at the moment.

It is also possible that the function of MeCP2-NCoR/SMRT interaction is not related to histone deacetylation, but affects other proteins. Acetylation has been shown to regulate localisation, DNA-binding and stability of various proteins (reviewed in: Kouzarides, 2000). HDACs are also able to deacetylate non-histone substrates and influence their function, for example HDAC3-mediated deacetylation of NFκB subunit causes its export from the nucleus (Chen et al., 2001a). The comparison of the acetylated proteins in *Mecp2* wild-type and knock-out brains could provide an insight on this hypothesis.

The simple model of MeCP2 function proposed by us offers hope for developing new strategies to try to cure at least some of the Rett syndrome patients. A screen for molecules capable of overriding the R306C mutation by allowing NCoR/SMRT to bind to mutated MeCP2 is one approach that could to be taken. Another possibility is to search for NCoR/SMRT mutants that could interact with altered MeCP2. Perhaps the manipulation of the components of NCoR/SMRT would be an easier target for potential gene therapy than MeCP2. I believe that this study provides avenues for new research towards future therapy of Rett syndrome.

References

Adams, V.H., McBryant, S.J., Wade, P. a, Woodcock, C.L., and Hansen, J.C. (2007). Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. The Journal of Biological Chemistry *282*, 15057–15064.

Adkins, N.L., and Georgel, P.T. (2011). MeCP2: structure and function. Biochemistry and Cell Biology = Biochimie Et Biologie Cellulaire *89*, 1–11.

Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H., and Cardoso, M.C. (2007). MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. Nucleic Acids Research *35*, 5402–5408.

Aid, T., Kazantseva, A., Piirsoo, M., Palm, K., and Timmusk, T. (2007). Mouse and rat BDNF gene structure and expression revisited. Journal of Neuroscience Research *85*, 525–535.

Allan, A.M., Liang, X., Luo, Y., Pak, C., Li, X., Szulwach, K.E., Chen, D., Jin, P., and Zhao, X. (2008). The loss of methyl-CpG binding protein 1 leads to autism-like behavioral deficits. Human Molecular Genetics *17*, 2047–2057.

Alland, L., Muhle, R., Hou, H., Potes, J., Chin, L., Schreiber-Agus, N., and DePinho, R.A. (1997). Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. Nature *387*, 49–55.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nature Genetics *23*, 185–188.

Antequera, F., Macleod, D., and Bird, a P. (1989). Specific protection of methylated CpGs in mammalian nuclei. Cell *58*, 509–517.

Antequera, F., Tamame, M., Villanueva, J.R., and Santos, T. (1984). DNA methylation in the fungi. The Journal of Biological Chemistry *259*, 8033–8036.

Argentaro, A., Sim, H., Kelly, S., Preiss, S., Clayton, A., Jans, D. a, and Harley, V.R. (2003). A SOX9 defect of calmodulin-dependent nuclear import in campomelic dysplasia/autosomal sex reversal. The Journal of Biological Chemistry *278*, 33839–33847.

Ballestar, E., Ropero, S., Alaminos, M., Armstrong, J., Setien, F., Agrelo, R., Fraga, M.F., Herranz, M., Avila, S., Pineda, M., et al. (2005). The impact of MECP2 mutations in the expression patterns of Rett syndrome patients. Human Genetics *116*, 91–104.

Ballestar, E., Yusufzai, T.M., and Wolffe, a P. (2000). Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. Biochemistry *39*, 7100–7106.

Barak, O., Lazzaro, M. a, Lane, W.S., Speicher, D.W., Picketts, D.J., and Shiekhattar, R. (2003). Isolation of human NURF: a regulator of Engrailed gene expression. The EMBO Journal *22*, 6089–6100.

Bartels, S.J.J., Spruijt, C.G., Brinkman, A.B., Jansen, P.W.T.C., Vermeulen, M., and Stunnenberg, H.G. (2011). A SILAC-based screen for Methyl-CpG binding proteins identifies RBP-J as a DNA methylation and sequence-specific binding protein. PloS One *6*, e25884.

Baudino, T. a, Kraichely, D.M., Jefcoat, S.C., Winchester, S.K., Partridge, N.C., and MacDonald, P.N. (1998). Isolation and characterization of a novel coactivator protein, NCoA-62, involved in vitamin D-mediated transcription. The Journal of Biological Chemistry *273*, 16434–16441.

Beard, C., Li, E., and Jaenisch, R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. Genes & Development *9*, 2325–2334.

Bedford, M.T., Chan, D.C., and Leder, P. (1997). FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. The EMBO Journal *16*, 2376–2383.

Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, C. a, and Zoghbi, H.Y. (2009). Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. Human Molecular Genetics *18*, 2431–2442.

Berger, J., Sansom, O., Clarke, A., and Bird, A. (2007). MBD2 is required for correct spatial gene expression in the gut. Molecular and Cellular Biology *27*, 4049–4057.

Bhaskara, S., Chyla, B.J., Amann, J.M., Knutson, S.K., Cortez, D., Sun, Z.-W., and Hiebert, S.W. (2008). Deletion of histone deacetylase 3 reveals critical roles in S phase progression and DNA damage control. Molecular Cell *30*, 61–72.

Bibel, M., Richter, J., Schrenk, K., Tucker, K.L., Staiger, V., Korte, M., Goetz, M., and Barde, Y.-A. (2004). Differentiation of mouse embryonic stem cells into a defined neuronal lineage. Nature Neuroscience *7*, 1003–1009.

Bienvenu, T., and Chelly, J. (2006). Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. Nature Reviews. Genetics *7*, 415–426.

Bird, A., Taggart, M., Frommer, M., Miller, O.J., and Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell *40*, 91–99.

Bird, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Research *8*, 1499–1504.

Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. Nature *321*, 209–213.

Bird, A.P. (1995). Gene number, noise reduction and biological complexity. Trends in Genetics: TIG *11*, 94–100.

Bird, A.P., and Taggart, M.H. (1980). Variable patterns of total DNA and rDNA methylation in animals. Nucleic Acids Research *8*, 1485–1497.

Bird, A.P., Taggart, M.H., and Smith, B.A. (1979). Methylated and unmethylated DNA compartments in the sea urchin genome. Cell *17*, 889–901.

Bischoff, F.R., and Görlich, D. (1997). RanBP1 is crucial for the release of RanGTP from importin β-related nuclear transport factors. FEBS Letters *419*, 249–254.

Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science (New York, N.Y.) *317*, 1760–1764.

Boyes, J., and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell *64*, 1123–1134.

Bracaglia, G., Conca, B., Bergo, A., Rusconi, L., Zhou, Z., Greenberg, M.E., Landsberger, N., Soddu, S., and Kilstrup-Nielsen, C. (2009). Methyl-CpG-binding protein 2 is phosphorylated by homeodomain-interacting protein kinase 2 and contributes to apoptosis. EMBO Reports *10*, 1327–1333.

Brockdorff, N. (2011). Chromosome silencing mechanisms in X-chromosome inactivation: unknown unknowns. Development (Cambridge, England) *138*, 5057–5065.

Buck-Koehntop, B. a, Martinez-Yamout, M. a, Dyson, H.J., and Wright, P.E. (2012). Kaiso uses all three zinc fingers and adjacent sequence motifs for high affinity binding to sequence-specific and methyl-CpG DNA targets. FEBS Letters *586*, 734–739.

Buckmaster, P.S., and Schwartzkroin, P. a (1995). Physiological and morphological heterogeneity of dentate gyrus-hilus interneurons in the gerbil hippocampus in vivo. The European Journal of Neuroscience *7*, 1393–1402.

Buschdorf, J., and Strätling, W. (2004). A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. Journal of Molecular Medicine *82*, 135–143.

Buschhausen, G., Graessmann, M., and Graessmann, A. (1985). Inhibition of herpes simplex thymidine kinase gene expression by DNA methylation is an indirect effect. Nucleic Acids Research *13*, 5503–5513.

Chae, J.H., Hwang, H., Hwang, Y.S., Cheong, H.J., and Kim, K.J. (2004). Influence of MECP2 gene mutation and X-chromosome inactivation on the Rett syndrome phenotype. Journal of Child Neurology *19*, 503–508.

Chahrour, M., Jung, S.Y., Shaw, C., Zhou, X., Wong, S.T.C., Qin, J., and Zoghbi, H.Y. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. Science (New York, N.Y.) *320*, 1224–1229.

Chandler, S.P., Guschin, D., Landsberger, N., and Wolffe, a P. (1999). The methyl-CpG binding transcriptional repressor MeCP2 stably associates with nucleosomal DNA. Biochemistry *38*, 7008–7018.

Chang, Q., Khare, G., Dani, V., Nelson, S., and Jaenisch, R. (2006). The disease progression of Mecp2 mutant mice is affected by the level of BDNF expression. Neuron *49*, 341–348.

Chao, H.-T., and Zoghbi, H.Y. (2012). MeCP2: only 100% will do. Nature Neuroscience *15*, 176–177.

Chao, L., Stratton, M., Lee, I., and Rosenberg, O. (2011). A Mechanism for Tunable Autoinhibition in the Structure of a Human Ca2+/Calmodulin-Dependent Kinase II Holoenzyme. Cell *146*, 732–745.

Chen, J.D., and Evans, R.M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature *377*, 454–457.

Chen, L., Fischle, W., Verdin, E., and Greene, W.C. (2001a). Duration of nuclear NF-kappaB action regulated by reversible acetylation. Science (New York, N.Y.) *293*, 1653–1657.

Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001b). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nature Genetics *27*, 327–331.

Chen, W.G., Chang, Q., Lin, Y., Meissner, A., West, A.E., Griffith, E.C., Jaenisch, R., and Greenberg, M.E. (2003). Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. Science (New York, N.Y.) *302*, 885–889.

Cheval, H., Guy, J., Merusi, C., De Sousa, D., Selfridge, J., and Bird, A. (2012). Postnatal inactivation reveals enhanced requirement for MeCP2 at distinct age windows. Human Molecular Genetics *21*, 3806–3814.

Christodoulou, J., Grimm, A., Maher, T., and Bennetts, B. (2003). RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. Human Mutation *21*, 466–472.

Clouaire, T., de Las Heras, J.I., Merusi, C., and Stancheva, I. (2010). Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA. Nucleic Acids Research *38*, 4620–4634.

Cohen, S., Gabel, H.W., Hemberg, M., Hutchinson, A.N., Sadacca, L.A., Ebert, D.H., Harmin, D. a, Greenberg, R.S., Verdine, V.K., Zhou, Z., et al. (2011). Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. Neuron *72*, 72–85.

Cohen-Cory, S., Kidane, A.H., Shirkey, N.J., and Marshak, S. (2010). Brain-derived neurotrophic factor and the development of structural neuronal connectivity. Developmental Neurobiology *70*, 271–288.

Colantuoni, C., Jeon, O.H., Hyder, K., Chenchik, a, Khimani, a H., Narayanan, V., Hoffman, E.P., Kaufmann, W.E., Naidu, S., and Pevsner, J. (2001). Gene expression profiling in postmortem Rett Syndrome brain: differential gene expression and patient classification. Neurobiology of Disease *8*, 847–865.

Collins, A.L., Levenson, J.M., Vilaythong, A.P., Richman, R., Armstrong, D.L., Noebels, J.L., David Sweatt, J., and Zoghbi, H.Y. (2004). Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. Human Molecular Genetics *13*, 2679–2689.

Coy, J.F., Sedlacek, Z., Bächner, D., Delius, H., and Poustka, A. (1999). A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3'-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. Human Molecular Genetics *8*, 1253–1262.

Cross, S.H., Meehan, R.R., Nan, X., and Bird, A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. Nature Genetics *16*, 256–259.

Daniel, J.M., Spring, C.M., Crawford, H.C., Reynolds, A.B., and Baig, A. (2002). The p120(ctn)-binding partner Kaiso is a bi-modal DNA-binding protein that recognizes both a sequence-specific consensus and methylated CpG dinucleotides. Nucleic Acids Research *30*, 2911–2919.

Delgado, I.J., Kim, D.S., Thatcher, K.N., LaSalle, J.M., and Van den Veyver, I.B. (2006). Expression profiling of clonal lymphocyte cell cultures from Rett syndrome patients. BMC Medical Genetics *7*, 61.

Deng, V., Matagne, V., Banine, F., Frerking, M., Ohliger, P., Budden, S., Pevsner, J., Dissen, G. a, Sherman, L.S., and Ojeda, S.R. (2007). FXYD1 is an MeCP2 target gene overexpressed in the brains of Rett syndrome patients and Mecp2-null mice. Human Molecular Genetics *16*, 640–650.

Derecki, N.C., Cronk, J.C., Lu, Z., Xu, E., Abbott, S.B.G., Guyenet, P.G., and Kipnis, J. (2012). Wild-type microglia arrest pathology in a mouse model of Rett syndrome. Nature *484*, 105–109.

Drewell, R. a, Goddard, C.J., Thomas, J.O., and Surani, M.A. (2002). Methylation-dependent silencing at the H19 imprinting control region by MeCP2. Nucleic Acids Research *30*, 1139–1144.

D'Esposito, M., Quaderi, N. a, Ciccodicola, a, Bruni, P., Esposito, T., D'Urso, M., and Brown, S.D. (1996). Isolation, physical mapping, and northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2. Mammalian Genome: Official Journal of the International Mammalian Genome Society *7*, 533–535.

Edelman, a M., Blumenthal, D.K., and Krebs, E.G. (1987). Protein serine/threonine kinases. Annual Review of Biochemistry *56*, 567–613.

Edwards, C.A., and Ferguson-Smith, A.C. (2007). Mechanisms regulating imprinted genes in clusters. Current Opinion in Cell Biology *19*, 281–289.

Ehrlich, M. (2009). DNA hypomethylation in cancer cells. Epigenomics *1*, 239–259.

Ehrlich, M., Wilson, G.G., Kuo, K.C., and Gehrke, C.W. (1987). N4-methylcytosine as a minor base in bacterial DNA. J. Bacteriol. *169*, 939–943.

El-Osta, a, and Wolffe, a P. (2001). Analysis of chromatin-immunopurified MeCP2-associated fragments. Biochemical and Biophysical Research Communications *289*, 733–737.

El-Osta, A., Kantharidis, P., Zalcberg, J.R., and Wolffe, A.P. (2002). Precipitous release of methyl-CpG binding protein 2 and histone deacetylase 1 from the methylated human multidrug resistance gene (MDR1) on activation. Molecular and Cellular Biology *22*, 1844–1857.

Van Esch, H., Bauters, M., Ignatius, J., Jansen, M., Raynaud, M., Hollanders, K., Lugtenberg, D., Bienvenu, T., Jensen, L.R., Gecz, J., et al. (2005). Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. American Journal of Human Genetics *77*, 442–453.

Esteller, M. (2007). Epigenetic gene silencing in cancer: the DNA hypermethylome. Human Molecular Genetics *16 Spec No*, R50–9.

Fasman, G.D. (1989). Prediction of Protein Structure and the Principles of Protein Conformation (Google eBook) (Springer).

Filion, G.G.J.P., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E., and Defossez, P.-A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. Molecular and Cellular Biology *26*, 169–181.

Forlani, G., Giarda, E., Ala, U., Di Cunto, F., Salani, M., Tupler, R., Kilstrup-Nielsen, C., and Landsberger, N. (2010). The MeCP2/YY1 interaction regulates ANT1 expression at 4q35: novel hints for Rett syndrome pathogenesis. Human Molecular Genetics *19*, 3114–3123.

Fraga, M.F. (2003). The affinity of different MBD proteins for a specific methylated locus depends on their intrinsic binding properties. Nucleic Acids Research *31*, 1765–1774.

Froger, A., and Hall, J.E. (2007). Transformation of plasmid DNA into E. coli using the heat shock method. Journal of Visualized Experiments: JoVE 253.

Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H., and Nakao, M. (1999). Methylation-Mediated Transcriptional Silencing in Euchromatin by Methyl-CpG Binding Protein MBD1 Isoforms. Molecular and Cellular Biology *19*, 6415–6426.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. The Journal of Biological Chemistry *278*, 4035–4040.

Gadalla, K.K.E., Bailey, M.E.S., and Cobb, S.R. (2011). MeCP2 and Rett syndrome: reversibility and potential avenues for therapy. The Biochemical Journal *439*, 1–14.

Galvão, T.C., and Thomas, J.O. (2005). Structure-specific binding of MeCP2 to four-way junction DNA through its methyl CpG-binding domain. Nucleic Acids Research *33*, 6603–6609.

Georgel, P.T., Horowitz-Scherer, R. a, Adkins, N., Woodcock, C.L., Wade, P. a, and Hansen, J.C. (2003). Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. The Journal of Biological Chemistry *278*, 32181–32188.

Ghosh, R.P., Horowitz-Scherer, R. a, Nikitina, T., Gierasch, L.M., and Woodcock, C.L. (2008). Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions. The Journal of Biological Chemistry *283*, 20523–20534.

Ghosh, R.P., Horowitz-Scherer, R. a, Nikitina, T., Shlyakhtenko, L.S., and Woodcock, C.L. (2010a). MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. Molecular and Cellular Biology *30*, 4656–4670.

Ghosh, R.P., Nikitina, T., Horowitz-Scherer, R. a, Gierasch, L.M., Uversky, V.N., Hite, K., Hansen, J.C., and Woodcock, C.L. (2010b). Unique physical properties and interactions of the domains of methylated DNA binding protein 2. Biochemistry *49*, 4395–4410.

Goffin, D., Allen, M., Zhang, L., Amorim, M., Wang, I.-T.J., Reyes, A.-R.S., Mercado-Berton, A., Ong, C., Cohen, S., Hu, L., et al. (2012). Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. Nature Neuroscience *15*, 274–283.

Goldfarb, D.S., Corbett, A.H., Mason, D.A., Harreman, M.T., and Adam, S. a (2004). Importin alpha: a multipurpose nuclear-transport receptor. Trends in Cell Biology *14*, 505–514.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annual Review of Biochemistry *74*, 481–514.

Gonzales, M.L., Adams, S., Dunaway, K.W., and Lasalle, J.M. (2012). Phosphorylation of Distinct Sites in MeCP2 Modifies Cofactor Associations and the Dynamics of Transcriptional Regulation. Molecular and Cellular Biology *32*, 2894–2903.

Gregory, R.I., Randall, T.E., Johnson, C.A., Khosla, S., Hatada, I., O'Neill, L.P., Turner, B.M., and Feil, R. (2001). DNA methylation is linked to deacetylation of histone H3, but not H4, on the imprinted genes Snrpn and U2af1-rs1. Molecular and Cellular Biology *21*, 5426–5436.

Guarda, A., Bolognese, F., Bonapace, I.M., and Badaracco, G. (2009). Interaction between the inner nuclear membrane lamin B receptor and the heterochromatic methyl binding protein, MeCP2. Experimental Cell Research *315*, 1895–1903.

Le Guezennec, X., Vermeulen, M., Brinkman, A.B., Hoeijmakers, W.A.M., Cohen, A., Lasonder, E., and Stunnenberg, H.G. (2006). MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Molecular and Cellular Biology *26*, 843–851.

Guo, J.U., Su, Y., Zhong, C., Ming, G., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423–434.

Guy, J., Cheval, H., Selfridge, J., and Bird, A. (2011). The role of MeCP2 in the brain. Annual Review of Cell and Developmental Biology *27*, 631–652.

Guy, J., Gan, J., Selfridge, J., Cobb, S., and Bird, A. (2007). Reversal of neurological defects in a mouse model of Rett syndrome. Science (New York, N.Y.) *315*, 1143–1147.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nature Genetics *27*, 322–326.

Hagberg, B., Aicardi, J., Dias, K., and Ramos, O. (1983). A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Annals of Neurology *14*, 471–479.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series *41*, 95–98.

Harikrishnan, K., Pal, S., Yarski, M., Baker, E.K., Chow, M.Z., de Silva, M.G., Okabe, J., Wang, L., Jones, P.L., Sif, S., et al. (2006). Reply to "Testing for association between MeCP2 and the brahma-associated SWI/SNF chromatin-remodeling complex."Nature Genetics *38*, 964–967.

Harikrishnan, K.N., Bayles, R., Ciccotosto, G.D., Maxwell, S., Cappai, R., Pelka, G.J., Tam, P.P.L., Christodoulou, J., and El-Osta, A. (2010). Alleviating transcriptional inhibition of the norepinephrine slc6a2 transporter gene in depolarized neurons. The Journal of Neuroscience: the Official Journal of the Society for Neuroscience *30*, 1494–1501.

Harikrishnan, K.N., Chow, M.Z., Baker, E.K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Craig, J.M., Jones, P.L., Sif, S., et al. (2005). Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. Nature Genetics *37*, 254–264.

Hark, a T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486–489.

Heard, E., and Disteche, C.M. (2006). Dosage compensation in mammals: fine-tuning the expression of the X chromosome. Genes & Development *20*, 1848–1867.

Heinzel, T., Lavinsky, R.M., Mullen, T.M., Söderstrom, M., Laherty, C.D., Torchia, J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R., et al. (1997). A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature *387*, 43–48.

Hendrich, B., Abbott, C., and McQueen, H. (1999a). Genomic structure and chromosomal mapping of the murine and human Mbd1, Mbd2, Mbd3, and Mbd4 genes. Mammalian Genome *912*, 906–912.

Hendrich, B., and Bird, A (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Molecular and Cellular Biology *18*, 6538–6547.

Hendrich, B., Guy, J., Ramsahoye, B., Wilson, V. a, and Bird, A (2001). Closely related proteins MBD2 and MBD3 play distinctive but interacting roles in mouse development. Genes & Development *15*, 710–723.

Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J., and Bird, A (1999b). The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature *401*, 301–304.

Hendrich, B., and Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. Trends in Genetics: TIG *19*, 269–277.

Ho, K.L., McNae, I.W., Schmiedeberg, L., Klose, R.J., Bird, A.P., and Walkinshaw, M.D. (2008). MeCP2 binding to DNA depends upon hydration at methyl-CpG. Molecular Cell *29*, 525–531.

Hoelz, A., Nairn, A.C., and Kuriyan, J. (2003). Crystal structure of a tetradecameric assembly of the association domain of Ca2+/calmodulin-dependent kinase II. Molecular Cell *11*, 1241–1251.

Horike, S., Cai, S., Miyano, M., Cheng, J.-F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nature Genetics *37*, 31–40.

Hu, K., Nan, X., Bird, A., and Wang, W. (2006). Testing for association between MeCP2 and the brahma-associated SWI/SNF chromatin-remodeling complex. Nature Genetics *38*, 962–4; author reply 964–7.

Hutchins, A.S., Mullen, A.C., Lee, H.W., Sykes, K.J., High, F.A., Hendrich, B.D., Bird, A.P., and Reiner, S.L. (2002). Gene silencing quantitatively controls the function of a developmental trans-activator. Molecular Cell *10*, 81–91.

Hörlein, A.J., Näär, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Söderström, M., and Glass, C.K. (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature *377*, 397–404.

Hörz, W., Altenburger, W., and Horz, W. (1981). Nucleotide sequence of mouse satellite DNA. Nucleic Acids Research *9*, 683–696.

Illingworth, R., Kerr, A., Desousa, D., Jørgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., et al. (2008). A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biology *6*, e22.

Illingworth, R.S., and Bird, A.P. (2009). CpG islands--'a rough guide'. FEBS Letters *583*, 1713–1720.

Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R., and Bird, A.P. (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. PLoS Genetics *6*, e1001134.

Imamura, T., Kerjean, A., Heams, T., Kupiec, J.-J., Thenevin, C., and Pàldi, A. (2005). Dynamic CpG and non-CpG methylation of the Peg1/Mest gene in the mouse oocyte and preimplantation embryo. The Journal of Biological Chemistry *280*, 20171–20175.

Inoue, A., and Zhang, Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science (New York, N.Y.) *334*, 194.

Ishibashi, T., Thambirajah, A. a, and Ausió, J. (2008). MeCP2 preferentially binds to methylated linker DNA in the absence of the terminal tail of histone H3 and independently of histone acetylation. FEBS Letters *582*, 1157–1162.

Ishihama, Y., Rappsilber, J., Andersen, J.S., and Mann, M. (2002). Microcolumns with self-assembled particle frits for proteomics. Journal of Chromatography. A *979*, 233–239.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science (New York, N.Y.) *333*, 1300–1303.

Jackson, I.J. (2001). Mouse genomics: making sense of the sequence. Current Biology: CB *11*, R311–4.

Jacobson, R.H., Zhang, X.J., DuBose, R.F., and Matthews, B.W. (1994). Three-dimensional structure of beta-galactosidase from E. coli. Nature *369*, 761–766.

Jeffery, L., and Nakielny, S. (2004). Components of the DNA methylation system of chromatin control are RNA-binding proteins. The Journal of Biological Chemistry *279*, 49479–49487.

Jepsen, K., Gleiberman, A.S., Shi, C., Simon, D.I., and Rosenfeld, M.G. (2008). Cooperative regulation in development by SMRT and FOXP1. 740–745.

Jepsen, K., Hermanson, O., Onami, T.M., Gleiberman, a S., Lunyak, V., McEvilly, R.J., Kurokawa, R., Kumar, V., Liu, F., Seto, E., et al. (2000). Combinatorial roles of the nuclear receptor corepressor in transcription and development. Cell *102*, 753–763.

Jepsen, K., Solum, D., Zhou, T., McEvilly, R.J., Kim, H.-J., Glass, C.K., Hermanson, O., and Rosenfeld, M.G. (2007). SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. Nature *450*, 415–419.

Jiang, Y., Matevossian, A., Huang, H.-S., Straubhaar, J., and Akbarian, S. (2008). Isolation of neuronal chromatin from brain tissue. BMC Neuroscience *9*, 42.

Jones, P.A., and Taylor, S.M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. Cell *20*, 85–93.

Jones, P.L., Sachs, L.M., Rouse, N., Wade, P. a, and Shi, Y.B. (2001). Multiple N-CoR complexes contain distinct histone deacetylases. The Journal of Biological Chemistry *276*, 8807–8811.

Jones, P.L., Veenstra, G.J., Wade, P. a, Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, a P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics *19*, 187–191.

Jordan, C., Li, H.H., Kwan, H.C., and Francke, U. (2007). Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets. BMC Medical Genetics *8*, 36.

Jørgensen, H.F., Ben-Porath, I., and Bird, A.P. (2004). Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. Molecular and Cellular Biology *24*, 3387–3395.

Kaji, K., Nichols, J., and Hendrich, B. (2007). Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. Development (Cambridge, England) *134*, 1123–1132.

Kaludov, N.K., and Wolffe, a P. (2000). MeCP2 driven transcriptional repression in vitro: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. Nucleic Acids Research *28*, 1921–1928.

Kang, M.-R., Lee, S.-W., Um, E., Kang, H.T., Hwang, E.S., Kim, E.-J., and Um, S.-J. (2010). Reciprocal roles of SIRT1 and SKIP in the regulation of RAR activity: implication in the retinoic acid-induced neuronal differentiation of P19 cells. Nucleic Acids Research *38*, 822–831.

Karagianni, P., and Wong, J. (2007). HDAC3: taking the SMRT-N-CoRrect road to repression. Oncogene *26*, 5439–5449.

Kass, S.U., Pruss, D., and Wolffe, A.P. (1997). How does DNA methylation repress transcription? Trends in Genetics *13*, 444–449.

Kernohan, K.D., Jiang, Y., Tremblay, D.C., Bonvissuto, A.C., Eubanks, J.H., Mann, M.R.W., and Bérubé, N.G. (2010). ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. Developmental Cell *18*, 191–202.

Kerr, B., Soto C, J., Saez, M., Abrams, A., Walz, K., and Young, J.I. (2012). Transgenic complementation of MeCP2 deficiency: phenotypic rescue of Mecp2-null mice by isoform-specific transgenes. European Journal of Human Genetics: EJHG *20*, 69–76.

Kersey, P.J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E., and Apweiler, R. (2004). The International Protein Index: an integrated database for proteomics experiments. Proteomics *4*, 1985–1988.

Kimura, H., and Shiota, K. (2003). Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. The Journal of Biological Chemistry *278*, 4806–4812.

Klein, M.E., Lioy, D.T., Ma, L., Impey, S., Mandel, G., and Goodman, R.H. (2007). Homeostatic regulation of MeCP2 expression by a CREB-induced microRNA. Nature Neuroscience *10*, 1513–1514.

Klose, R.J., and Bird, A.P. (2004). MeCP2 behaves as an elongated monomer that does not stably associate with the Sin3a chromatin remodeling complex. The Journal of Biological Chemistry *279*, 46490–46496.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends in Biochemical Sciences *31*, 89–97.

Klose, R.J., Sarraf, S. a, Schmiedeberg, L., McDermott, S.M., Stancheva, I., and Bird, A.P. (2005). DNA binding selectivity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. Molecular Cell *19*, 667–678.

Kohno, K., Izumi, H., Uchiumi, T., Ashizuka, M., and Kuwano, M. (2003). The pleiotropic functions of the Y-box-binding protein, YB-1. BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology *25*, 691–698.

Kokura, K., Kaul, S.C., Wadhwa, R., Nomura, T., Khan, M.M., Shinagawa, T., Yasukawa, T., Colmenares, C., and Ishii, S. (2001). The Ski protein family is required for MeCP2-mediated transcriptional repression. The Journal of Biological Chemistry *276*, 34115–34121.

Kondo, E., Gu, Z., Horii, A., and Fukushige, S. (2005). The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16INK4a and hMLH1 genes. Molecular and Cellular Biology *25*, 4388–4396.

Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., and Yanagawa, H. (2009). Six classes of nuclear localization signals specific to different binding grooves of importin alpha. The Journal of Biological Chemistry *284*, 478–485.

Kouzarides, T. (2000). Acetylation: a regulatory modification to rival phosphorylation? The EMBO Journal *19*, 1176–1179.

Kriaucionis, S., and Bird, A. (2003). DNA methylation and Rett syndrome. Human Molecular Genetics *12 Spec No*, R221–7.

Kriaucionis, S., and Bird, A. (2004). The major form of MeCP2 has a novel N-terminus generated by alternative splicing. Nucleic Acids Research *32*, 1818–1823.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science (New York, N.Y.) *324*, 929–930.

Kriaucionis, S., Paterson, A., Curtis, J., Guy, J., Macleod, N., and Bird, A. (2006). Gene expression analysis exposes mitochondrial abnormalities in a mouse model of Rett syndrome. Molecular and Cellular Biology *26*, 5033–5042.

Krithivas, A., Fujimuro, M., Weidner, M., Young, D.B., Hayward, S.D., Verma, S.C., Borah, S., and Robertson, E.S. (2004). Protein Interactions Targeting the Latency-Associated Nuclear Antigen of Kaposi's Sarcoma-Associated Herpesvirus to Cell Chromosomes. Journal of Virology *78*, 10348–10359.

Kudo, S. (1998). Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated. Molecular and Cellular Biology *18*, 5492–5499.

Kudo, S., Nomura, Y., Segawa, M., Fujita, N., Nakao, M., Dragich, J., Schanen, C., and Tamura, M. (2001). Functional analyses of MeCP2 mutations associated with Rett syndrome using transient expression systems. Brain & Development *23 Suppl 1*, S165–73.

Kudo, S., Nomura, Y., Segawa, M., Fujita, N., Nakao, M., Schanen, C., and Tamura, M. (2003). Heterogeneity in residual function of MeCP2 carrying missense mutations in the methyl CpG binding domain. Journal of Medical Genetics *40*, 487–493.

Kumar, A., Kamboj, S., Malone, B.M., Kudo, S., Twiss, J.L., Czymmek, K.J., LaSalle, J.M., and Schanen, N.C. (2008). Analysis of protein domains and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 in vivo. Journal of Cell Science *121*, 1128–1137.

Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R., and Görlich, D. (1997). Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. Cell *90*, 1061–1071.

Köhler, M., Speck, C., Christiansen, M., Bischoff, F.R., Prehn, S., Haller, H., Görlich, D., and Hartmann, E. (1999). Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import. Molecular and Cellular Biology *19*, 7782–7791.

LaSalle, J.M., Goldstine, J., Balmer, D., and Greco, C.M. (2001). Quantitative localization of heterogeneous methyl-CpG-binding protein 2 (MeCP2) expression phenotypes in normal and Rett syndrome brain by laser scanning cytometry. Human Molecular Genetics *10*, 1729–1740.

Laengle-Rouault, F., Maenhaut-Michel, G., and Radman, M. (1986). GATC sequence and mismatch repair in Escherichia coli. The EMBO Journal *5*, 2009–2013.

Laget, S., Joulie, M., Le Masson, F., Sasai, N., Christians, E., Pradhan, S., Roberts, R.J., and Defossez, P.-A. (2010). The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. PloS One *5*, e11982.

Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E., and Corbett, A.H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. The Journal of Biological Chemistry *282*, 5101–5105.

Larimore, J.L., Chapleau, C. a, Kudo, S., Theibert, A., Percy, A.K., and Pozzo-Miller, L. (2009). Bdnf overexpression in hippocampal neurons prevents dendritic atrophy caused by Rett-associated MECP2 mutations. Neurobiology of Disease *34*, 199–211.

Leong, G.M., Subramaniam, N., Issa, L.L., Barry, J.B., Kino, T., Driggers, P.H., Hayman, M.J., Eisman, J.A., and Gardiner, E.M. (2004). Ski-interacting protein, a bifunctional nuclear receptor coregulator that interacts with N-CoR/SMRT and p300. Biochemical and Biophysical Research Communications *315*, 1070–1076.

Levy, D.L., and Heald, R. (2010). Nuclear size is regulated by importin α and Ntf2 in Xenopus. Cell *143*, 288–298.

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, a (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell *69*, 905–914.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362–365.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, J., Wang, J., Nawaz, Z., Liu, J.M., Qin, J., and Wong, J. (2000). Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. The EMBO Journal *19*, 4342–4350.

Lieberman, D.N., and Mody, I. (1999). Casein kinase-II regulates NMDA channel function in hippocampal neurons. Nature Neuroscience *2*, 125–132.

Lindsay, M.E., Plafker, K., Smith, A.E., Clurman, B.E., and Macara, I.G. (2002). Npap60/Nup50 is a tri-stable switch that stimulates importin-alpha:beta-mediated nuclear protein import. Cell *110*, 349–360.

Lioy, D.T., Garg, S.K., Monaghan, C.E., Raber, J., Foust, K.D., Kaspar, B.K., Hirrlinger, P.G., Kirchhoff, F., Bissonnette, J.M., Ballas, N., et al. (2011). A role for glia in the progression of Rett's syndrome. Nature *475*, 497–500.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Llorian, M., Beullens, M., Andrés, I., Ortiz, J.-M., and Bollen, M. (2004). SIPP1, a novel pre-mRNA splicing factor and interactor of protein phosphatase-1. The Biochemical Journal *378*, 229–238.

Lorincz, M.M.C., Schübeler, D., and Groudine, M. (2001). Methylation-Mediated Proviral Silencing Is Associated with MeCP2 Recruitment and Localized Histone H3 Deacetylation. Molecular and Cellular Biology *21*, 7913–7922.

Lucchesi, W., Mizuno, K., and Giese, K.P. (2011). Novel insights into CaMKII function and regulation during memory formation. Brain Research Bulletin *85*, 2–8.

Luikenhuis, S., Giacometti, E., Beard, C.F., and Jaenisch, R. (2004). Expression of MeCP2 in postmitotic neurons rescues Rett syndrome in mice. Proceedings of the National Academy of Sciences of the United States of America *101*, 6033–6038.

Lunyak, V.V., Burgess, R., Prefontaine, G.G., Nelson, C., Sze, S.-H., Chenoweth, J., Schwartz, P., Pevzner, P. a, Glass, C., Mandel, G., et al. (2002). Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. Science (New York, N.Y.) *298*, 1747–1752.

Lyko, F., Ramsahoye, B.H., and Jaenisch, R. (2000). DNA methylation in Drosophila melanogaster. Nature *408*, 538–540.

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372–373.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. The Journal of Biological Chemistry *286*, 35334–35338.

Marchetto, M.C.N., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell *143*, 527–539.

Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R., Scala, E., Longo, I., Grosso, S., Pescucci, C., et al. (2005). CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. Human Molecular Genetics *14*, 1935–1946.

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y.E. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. Science (New York, N.Y.) *302*, 890–893.

Martín Caballero, I., Hansen, J., Leaford, D., Pollard, S., and Hendrich, B.D. (2009). The methyl-CpG binding proteins Mecp2, Mbd2 and Kaiso are dispensable for mouse embryogenesis, but play a redundant function in neural differentiation. PloS One *4*, e4315.

Mason, D.A., Máthé, E., Mathe, E., Fleming, R.J., and Goldfarb, D.S. (2003). The Drosophila melanogaster importin {alpha}3 Locus Encodes an Essential Gene Required for the Development of Both Larval and Adult Tissues. Genetics *165*, 1943–1958.

Matarazzo, M.R., De Bonis, M.L., Strazzullo, M., Cerase, A., Ferraro, M., Vastarelli, P., Ballestar, E., Esteller, M., Kudo, S., and D'Esposito, M. (2007). Multiple binding of methyl-CpG and polycomb proteins in long-term gene silencing events. Journal of Cellular Physiology *210*, 711–719.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature *466*, 253–257.

McDonald, L.E., Paterson, C. a, and Kay, G.F. (1998). Bisulfite genomic sequencing-derived methylation profile of the xist gene throughout early mouse development. Genomics *54*, 379–386.

Meehan, R.R., Lewis, J.D., and Bird, a P. (1992). Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Research *20*, 5085–5092.

Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L., and Bird, a P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell *58*, 499–507.

Meggio, F., and Pinna, L. a (2003). One-thousand-and-one substrates of protein kinase CK2? FASEB Journal *17*, 349–368.

Melchior, F. (2000). SUMO--nonclassical ubiquitin. Annual Review of Cell and Developmental Biology *16*, 591–626.

Millar, C.B., Guy, J., Sansom, O.J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P.D., Bishop, S.M., Clarke, A.R., and Bird, A. (2002). Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. Science (New York, N.Y.) *297*, 403–405.

Miller, O.J., Schnedl, W., Allen, J., and Erlanger, B.F. (1974). 5-Methylcytosine localised in mammalian constitutive heterochromatin. Nature *251*, 636–637.

Miyake, K., and Nagai, K. (2007). Phosphorylation of methyl-CpG binding protein 2 (MeCP2) regulates the intracellular localization during neuronal cell differentiation. Neurochemistry International *50*, 264–270.

Mnatzakanian, G.N., Lohi, H., Munteanu, I., Alfred, S.E., Yamada, T., MacLeod, P.J.M., Jones, J.R., Scherer, S.W., Schanen, N.C., Friez, M.J., et al. (2004). A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. Nature Genetics *36*, 339–341.

Mohandas, T., Sparkes, R.S., and Shapiro, L.J. (1981). Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. Science (New York, N.Y.) *211*, 393–396.

Muotri, A.R., Marchetto, M.C.N., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. Nature *468*, 443–446.

Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nature Chemical Biology *8*, 751-758.

Nakai, K., and Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends in Biochemical Sciences *24*, 34–36.

Nan, X., Campoy, F.J., and Bird, a (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471–481.

Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M.J., Shu, X., Kriaucionis, S., and Bird, A. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. Proceedings of the National Academy of Sciences of the United States of America *104*, 2709–2714.

Nan, X., Meehan, R.R., and Bird, a (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Research *21*, 4886–4892.

Nan, X., Ng, H.H., Johnson, C. a, Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, a (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386–389.

Nan, X., Tate, P., Li, E., and Bird, a (1996). DNA methylation specifies chromosomal localization of MeCP2. Molecular and Cellular Biology *16*, 414–421.

Nectoux, J., Fichou, Y., Rosas-Vargas, H., Cagnard, N., Bahi-Buisson, N., Nusbaum, P., Letourneur, F., Chelly, J., and Bienvenu, T. (2010). Cell cloning-based transcriptome analysis in Rett patients: relevance to the pathogenesis of Rett syndrome of new human MeCP2 target genes. Journal of Cellular and Molecular Medicine *14*, 1962–1974.

Ng, H.-H.H., Jeppesen, P., and Bird, A. (2000). Active Repression of Methylated Genes by the Chromosomal Protein MBD1. Molecular and Cellular Biology *20*, 1394–1406.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C. a, Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, a (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nature Genetics *23*, 58–61.

Nguyen, C.T., Gonzales, F. a, and Jones, P. a (2001). Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: correlation of accessibility, methylation, MeCP2 binding and acetylation. Nucleic Acids Research *29*, 4598–4606.

Nicolas, E., Morales, V., Magnaghi-Jaulin, L., Harel-Bellan, a, Richard-Foy, H., and Trouche, D. (2000). RbAp48 belongs to the histone deacetylase complex that associates with the retinoblastoma protein. The Journal of Biological Chemistry *275*, 9797–9804.

Nigg, E.A. (1997). Nucleocytoplasmic transport: signals, mechanisms and regulation. Nature *386*, 779–787.

Nikitina, T., Ghosh, R.P., Horowitz-Scherer, R. a, Hansen, J.C., Grigoryev, S. a, and Woodcock, C.L. (2007a). MeCP2-chromatin interactions include the formation of chromatosome-like structures and are altered in mutations causing Rett syndrome. The Journal of Biological Chemistry *282*, 28237–28245.

Nikitina, T., Shi, X., Ghosh, R.P., Horowitz-Scherer, R. a, Hansen, J.C., and Woodcock, C.L. (2007b). Multiple modes of interaction between the methylated DNA binding protein MeCP2 and chromatin. Molecular and Cellular Biology *27*, 864–877.

Nuber, U. a, Kriaucionis, S., Roloff, T.C., Guy, J., Selfridge, J., Steinhoff, C., Schulz, R., Lipkowitz, B., Ropers, H.H., Holmes, M.C., et al. (2005). Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome. Human Molecular Genetics *14*, 2247–2256.

Oberoi, J., Fairall, L., Watson, P.J., Yang, J.-C., Czimmerer, Z., Kampmann, T., Goult, B.T., Greenwood, J. a, Gooch, J.T., Kallenberger, B.C., et al. (2011). Structural basis for the assembly of the SMRT/NCoR core transcriptional repression machinery. Nature Structural & Molecular Biology *18*, 177–184.

Okano, M., Bell, D.W., Haber, D. a, and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A, Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Current Biology: CB *10*, 475–478.

Otis, K.O., Thompson, K.R., and Martin, K.C. (2006). Importin-mediated nuclear transport in neurons. Current Opinion in Neurobiology *16*, 329–335.

Paine, P.L., Moore, L.C., and Horowitz, S.B. (1975). Nuclear envelope permeability. Nature *254*, 109–114.

Palmer, B.R., and Marinus, M.G. (1994). The dam and dcm strains of Escherichia coli — a review. Gene *143*, 1–12.

Pastor, W. a, Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., et al. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature *473*, 394–397.

Pearson, E.C., Bates, D.L., Prospero, T.D., and Thomas, J.O. (1984). Neuronal nuclei and glial nuclei from mammalian cerebral cortex. Nucleosome repeat lengths, DNA contents and H1 contents. European Journal of Biochemistry / FEBS *144*, 353–360.

Pinna, L. a. (2002). Protein kinase CK2: a challenge to canons. Journal of Cell Science *115*, 3873–3878.

Prendergast, G.C., Ziff, E.B., and Bard, A.J. (1991). Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. Science (New York, N.Y.) *251*, 186–189.

Proffitt, J.H., Davie, J.R., Swinton, D., and Hattman, S. (1984). 5-Methylcytosine is not detectable in Saccharomyces cerevisiae DNA. Molecular and Cellular Biology *4*, 985–988.

Prokhortchouk, A., Hendrich, B., Jørgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes & Development *15*, 1613–1618.

Prokhortchouk, A., Sansom, O., Selfridge, J., Caballero, I.M., Salozhin, S., Aithozhina, D., Cerchietti, L., Meng, F.G., Augenlicht, L.H., Mariadason, J.M., et al. (2006). Kaiso-deficient mice show resistance to intestinal cancer. Molecular and Cellular Biology *26*, 199–208.

Pruschy, M., Ju, Y., Spitz, L., Carafoli, E., and Goldfarb, D.S. (1994). Facilitated nuclear transport of calmodulin in tissue culture cells. The Journal of Cell Biology *127*, 1527–1536.

Quensel, C., Friedrich, B., Sommer, T., Hartmann, E., and Kohler, M. (2004). In vivo analysis of importin alpha proteins reveals cellular proliferation inhibition and substrate specificity. Molecular and Cellular Biology *24*, 10246–10255.

Rabinowicz, P.D., Palmer, L.E., May, B.P., Hemann, M.T., Lowe, S.W., McCombie, W.R., and Martienssen, R. A. (2003). Genes and transposons are differentially methylated in plants, but not in mammals. Genome Research *13*, 2658–2664.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, a P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proceedings of the National Academy of Sciences of the United States of America *97*, 5237–5242.

Rangwala, S.H., and Richards, E.J. (2004). The value-added genome: building and maintaining genomic cytosine methylation landscapes. Current Opinion in Genetics & Development *14*, 686–691.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Analytical Chemistry *75*, 663–670.

Rebhan, M., Chalifa-Caspi, V., Prilusky, J., and Lancet, D. (1997). GeneCards: integrating information about genes, proteins and diseases. Trends in Genetics: TIG *13*, 163.

Rebhan, M., Chalifa-Caspi, V., Prilusky, J., and Lancet, D. (1998). GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. Bioinformatics (Oxford, England) *14*, 656–664.

Rechsteiner, M., and Rogers, S.W. (1996). PEST sequences and regulation by proteolysis. Trends in Biochemical Sciences *21*, 267–271.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science (New York, N.Y.) *293*, 1089–1093.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408–1419.

Rietveld, L.E.G., Caldenhoven, E., and Stunnenberg, H.G. (2002). In vivo repression of an erythroid-specific gene by distinct corepressor complexes. The EMBO Journal *21*, 1389–1397.

Robbins, J., Dilworth, S.M., Laskey, R.A., and Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell *64*, 615–623.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science (New York, N.Y.) *234*, 364–368.

Roloff, T.C., Ropers, H.H., and Nuber, U.A. (2003). Comparative study of methyl-CpG-binding domain proteins. BMC Genomics *4*, 1.

Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C., and Leonhardt, H. (2008). A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Molecular & Cellular Proteomics: MCP *7*, 282–289.

Ruzov, A., Dunican, D.S., Prokhortchouk, A., Pennings, S., Stancheva, I., Prokhortchouk, E., and Meehan, R.R. (2004). Kaiso is a genome-wide repressor of transcription that is essential for amphibian development. Development (Cambridge, England) *131*, 6185–6194.

Sakashita, E., Tatsumi, S., Werner, D., Endo, H., and Mayeda, A. (2004). Human RNPS1 and Its Associated Factors: a Versatile Alternative Pre-mRNA Splicing Regulator In Vivo. Molecular and Cellular Biology *24*, 1174–1187.

Samaco, R.C., Fryer, J.D., Ren, J., Fyffe, S., Chao, H.-T., Sun, Y., Greer, J.J., Zoghbi, H.Y., and Neul, J.L. (2008). A partial loss of function allele of methyl-CpG-binding protein 2 predicts a human neurodevelopmental syndrome. Human Molecular Genetics *17*, 1718–1727.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Developmental Biology *241*, 172–182.

Sauvé, F., McBroom, L.D., Gallant, J., Moraitis, A.N., Labrie, F., and Giguère, V. (2001). CIA, a novel estrogen receptor coactivator with a bifunctional nuclear receptor interacting determinant. Molecular and Cellular Biology *21*, 343–353.

Schmid, R., Tsujimoto, N., Qu, Q., Lei, H., and Li, E. (2008). A methyl-CpG-binding protein 2-enhanced green fluorescent protein reporter mouse model provides a new tool for studying the neuronal basis of Rett syndrome. Neuroreport *19*, 393–398.

Schmidt, D., and Schreiber, S. (1999). Molecular association between ATR and two components of the nucleosome remodeling and deacetylating complex, HDAC2 and CHD4. Biochemistry *38*, 14711–14717.

Schmiedeberg, L., Skene, P., Deaton, A., and Bird, A. (2009). A temporal threshold for formaldehyde crosslinking and fixation. PloS One *4*, e4636.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature Methods *9*, 671–675.

Schüle, B., Li, H.H., Fisch-Kohl, C., Purmann, C., and Francke, U. (2007). DLX5 and DLX6 expression is biallelic and not modulated by MeCP2 deficiency. American Journal of Human Genetics *81*, 492–506.

Selker, E.U., Tountas, N.A., Cross, S.H., Margolin, B.S., Murphy, J.G., Bird, A.P., and Freitag, M. (2003). The methylated component of the Neurospora crassa genome. Nature *422*, 893–897.

Shahbazian, M., Young, J., Yuva-Paylor, L., Spencer, C., Antalffy, B., Noebels, J., Armstrong, D., Paylor, R., and Zoghbi, H. (2002a). Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. Neuron *35*, 243–254.

Shahbazian, M.D., Antalffy, B., Armstrong, D.L., and Zoghbi, H.Y. (2002b). Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. Human Molecular Genetics *11*, 115–124.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. a, Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908–912.

Shen, X., Yu, L., Weir, J.W., and Gorovsky, M.A. (1995). Linker histones are not essential and affect chromatin condensation in vivo. Cell *82*, 47–56.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Analytical Chemistry *68*, 850–858.

Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.Z., and Cedar, H. (1999). DNA methylation represses transcription in vivo. Nature Genetics *22*, 203–206.

Sigrist, C.J. a, Cerutti, L., de Castro, E., Langendijk-Genevaux, P.S., Bulliard, V., Bairoch, A., and Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Research *38*, D161–6.

Simmen, M.W., Leitgeb, S., Charlton, J., Jones, S.J., Harris, B.R., Clark, V.H., and Bird, A. (1999). Nonmethylated transposable elements and methylated genes in a chordate genome. Science (New York, N.Y.) *283*, 1164–1167.

Singh, J., Saxena, A., Christodoulou, J., and Ravine, D. (2008). MECP2 genomic structure and function: insights from ENCODE. Nucleic Acids Research *36*, 6035–6047.

Sirianni, N., Naidu, S., Pereira, J., Pillotto, R.F., and Hoffman, E.P. (1998). Rett syndrome: confirmation of X-linked dominant inheritance, and localization of the gene to Xq28. American Journal of Human Genetics *63*, 1552–1558.

Skene, P., Illingworth, R., Webb, S., and Kerr, A. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Molecular Cell 457–468.

Stancheva, I., Collins, A.L., Van den Veyver, I.B., Zoghbi, H., and Meehan, R.R. (2003). A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. Molecular Cell *12*, 425–435.

Stein, R., Razin, A., and Cedar, H. (1982). In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. Proceedings of the National Academy of Sciences of the United States of America *79*, 3418–3422.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biology *12*, R54.

Sun, J.-M., Chen, H.Y., and Davie, J.R. (2007). Differential distribution of unmodified and phosphorylated histone deacetylase 2 in chromatin. The Journal of Biological Chemistry *282*, 33227–33236.

Suzuki, M., Yamada, T., Kihara-Negishi, F., Sakurai, T., and Oikawa, T. (2003). Direct association between PU.1 and MeCP2 that recruits mSin3A-HDAC complex for PU.1-mediated transcriptional repression. Oncogene *22*, 8688–8698.

Suzuki, M.M., Kerr, A.R.W., De Sousa, D., and Bird, A. (2007). CpG methylation is targeted to transcription units in an invertebrate genome. Genome Research *17*, 625–631.

Szulwach, K.E., Li, X., Smrt, R.D., Li, Y., Luo, Y., Lin, L., Santistevan, N.J., Li, W., Zhao, X., and Jin, P. (2010). Cross talk between microRNA and epigenetic regulation in adult neurogenesis. The Journal of Cell Biology *189*, 127–141.

Tacke, R., Tohyama, M., Ogawa, S., and Manley, J.L. (1998). Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. Cell *93*, 139–148.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W. a, Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science (New York, N.Y.) *324*, 930–935.

Tao, J., Hu, K., Chang, Q., Wu, H., Sherman, N.E., Martinowich, K., Klose, R.J., Schanen, C., Jaenisch, R., Wang, W., et al. (2009). Phosphorylation of MeCP2 at Serine 80 regulates its chromatin association and neurological function. Proceedings of the National Academy of Sciences of the United States of America *106*, 4882–4887.

Tapia, V.E., Nicolaescu, E., McDonald, C.B., Musi, V., Oka, T., Inayoshi, Y., Satteson, A.C., Mazack, V., Humbert, J., Gaffney, C.J., et al. (2010). Y65C missense mutation in the WW domain of the Golabi-Ito-Hall syndrome protein PQBP1 affects its binding activity and deregulates pre-mRNA splicing. The Journal of Biological Chemistry *285*, 19391–19401.

Tate, P., Skarnes, W., and Bird, A. (1996). The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. Nature Genetics *12*, 205–208.

Temudo, T., Santos, M., Ramos, E., Dias, K., Vieira, J.P., Moreira, A., Calado, E., Carrilho, I., Oliveira, G., Levy, A., et al. (2011). Rett syndrome with and without detected MECP2 mutations: an attempt to redefine phenotypes. Brain & Development *33*, 69–76.

Thambirajah, A. a, Eubanks, J.H., and Ausió, J. (2009). MeCP2 post-translational regulation through PEST domains: two novel hypotheses: potential relevance and implications for Rett syndrome. BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology *31*, 561–569.

Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R.W., Deaton, A., Andrews, R., James, K.D., et al. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature *464*, 1082–1086.

Traynor, J., Agarwal, P., Lazzeroni, L., and Francke, U. (2002). Gene expression patterns vary in clonal cell cultures from Rett syndrome females with eight different MECP2 mutations. BMC Medical Genetics *3*, 12.

Tudor, M., Akbarian, S., Chen, R.Z., and Jaenisch, R. (2002). Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. Proceedings of the National Academy of Sciences of the United States of America *99*, 15536–15541.

Tweedie, S., and Charlton, J. (1997). Methylation of genomes and genes at the invertebrate-vertebrate boundary. Molecular and Cellular Biology *17*, 1469–1475.

Urdinguio, R.G., Fernandez, A.F., Lopez-Nieva, P., Rossi, S., Huertas, D., Kulis, M., Liu, C.-G., Croce, C.M., Calin, G. a, and Esteller, M. (2010). Disrupted microRNA expression caused by Mecp2 loss in a mouse model of Rett syndrome. Epigenetics: Official Journal of the DNA Methylation Society *5*, 656–663.

Urdinguio, R.G., Lopez-Serra, L., Lopez-Nieva, P., Alaminos, M., Diaz-Uriarte, R., Fernandez, A.F., and Esteller, M. (2008). Mecp2-null mice provide new neuronal targets for Rett syndrome. PloS One *3*, e3669.

Venolia, L., Gartler, S.M., Wassman, E.R., Yen, P., Mohandas, T., and Shapiro, L.J. (1982). Transformation with DNA from 5-azacytidine-reactivated X chromosomes. Proceedings of the National Academy of Sciences of the United States of America *79*, 2352–2354.

Wada, R., Akiyama, Y., Hashimoto, Y., Fukamachi, H., and Yuasa, Y. (2010). miR-212 is downregulated and suppresses methyl-CpG-binding protein MeCP2 in human gastric cancer. International Journal of Cancer. Journal International Du Cancer *127*, 1106–1114.

Wagstaff, K.M., and Jans, D.A. (2009). Importins and Beyond: Non-Conventional Nuclear Transport Mechanisms. Traffic 1188–1198.

Wakefield, R.I., Smith, B.O., Nan, X., Free, a, Soteriou, a, Uhrin, D., Bird, a P., and Barlow, P.N. (1999). The solution structure of the domain from MeCP2 that binds to methylated DNA. Journal of Molecular Biology *291*, 1055–1065.

Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S., and Liu, J.M. (1998). ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proceedings of the National Academy of Sciences of the United States of America *95*, 10860–10865.

Wang, R., and Brattain, M.G. (2007). The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa. FEBS Letters *581*, 3164–3170.

Waterhouse, A.M., Procter, J.B., Martin, D.M. a, Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics (Oxford, England) *25*, 1189–1191.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737–738.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nature Genetics *39*, 457–466.

Weis, K. (2003). Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. Cell *112*, 441–451.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P. a C., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343–348.

Wilson, G., and Murray, N. (1991). Restriction and modification systems. Annual Review of Genetics *25*, 585–627.

Woodcock, C.L. (2006). Chromatin architecture. Current Opinion in Structural Biology *16*, 213–220.

Wu, H., Tao, J., Chen, P.J., Shahab, A., Ge, W., Hart, R.P., Ruan, X., Ruan, Y., and Sun, Y.E. (2010). Genome-wide analysis reveals methyl-CpG-binding protein 2-dependent regulation of microRNAs in a mouse model of Rett syndrome. Proceedings of the National Academy of Sciences of the United States of America *107*, 18161–18166.

Wyatt, G.R., and Cohen, S.S. (1952). A new pyrimidine base from bacteriophage nucleic acids. Nature *170*, 1072–1073.

Wyatt, G.R., and Cohen, S.S. (1953). The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. The Biochemical Journal *55*, 774–782.

Xiang, F., Zhang, Z., Clarke, a, Joseluiz, P., Sakkubai, N., Sarojini, B., Delozier-Blanchet, C.D., Hansmann, I., Edström, L., and Anvret, M. (1998). Chromosome mapping of Rett syndrome: a likely candidate region on the telomere of Xq. Journal of Medical Genetics *35*, 297–300.

Yamauchi, T. (2005). Neuronal Ca2+/calmodulin-dependent protein kinase II--discovery, progress in a quarter of a century, and perspective: implication for learning and memory. Biological & Pharmaceutical Bulletin *28*, 1342–1354.

Yang, C., van der Woerd, M.J., Muthurajan, U.M., Hansen, J.C., and Luger, K. (2011). Biophysical analysis and small-angle X-ray scattering-derived structures of MeCP2-nucleosome complexes. Nucleic Acids Research *39*, 4122–4135.

Yasui, D.H., Peddada, S., Bieda, M.C., Vallero, R.O., Hogart, A., Nagarajan, R.P., Thatcher, K.N., Farnham, P.J., and Lasalle, J.M. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. Proceedings of the National Academy of Sciences of the United States of America *104*, 19416–19421.

Yazdani, M., Deogracias, R., Guy, J., Poot, R. a, Bird, A., and Barde, Y.-A. (2012). Disease Modeling Using Embryonic Stem Cells: MeCP2 Regulates Nuclear Size and RNA Synthesis in Neurons. Stem Cells (Dayton, Ohio).

Yildirim, O., Li, R., Hung, J.-H., Chen, P.B., Dong, X., Ee, L.-S., Weng, Z., Rando, O.J., and Fazzio, T.G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell *147*, 1498–1510.

Yoder, J., Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends in Genetics: TIG *13*, 335–340.

Yoon, H.-G., Chan, D.W., Reynolds, A.B., Qin, J., and Wong, J. (2003). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. Molecular Cell *12*, 723–734.

Yoshida, M., Kijima, M., Akita, M., and Beppu, T. (2007). Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. The Journal of Biological Chemistry *52*, 1788–1789.

Young, J.I., Hong, E.P., Castle, J.C., Crespo-Barreto, J., Bowman, A.B., Rose, M.F., Kang, D., Richman, R., Johnson, J.M., Berget, S., et al. (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. Proceedings of the National Academy of Sciences of the United States of America *102*, 17551–17558.

Yu, F., Thiesen, J., and Strätling, W.H. (2000). Histone deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2. Nucleic Acids Research *28*, 2201–2206.

Yu, F., Zingler, N., Schumann, G., and Strätling, W.H. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Research *29*, 4493–4501.

Yu, I.J., Spector, D.L., Bae, Y.S., and Marshak, D.R. (1991). Immunocytochemical localization of casein kinase II during interphase and mitosis. The Journal of Cell Biology *114*, 1217–1232.

Yusufzai, T.M., and Wolffe, a P. (2000). Functional consequences of Rett syndrome mutations on human MeCP2. Nucleic Acids Research *28*, 4172–4179.

Zhang, J., Kalkum, M., Chait, B.T., and Roeder, R.G. (2002). The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. Molecular Cell *9*, 611–623.

Zhao, X., Ueba, T., Christie, B.R., Barkho, B., McConnell, M.J., Nakashima, K., Lein, E.S., Eadie, B.D., Willhoite, A.R., Muotri, A.R., et al. (2003). Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. Proceedings of the National Academy of Sciences of the United States of America *100*, 6777–6782.

Zhou, Z., Hong, E.J., Cohen, S., Zhao, W.-N., Ho, H.-Y.H., Schmidt, L., Chen, W.G., Lin, Y., Savner, E., Griffith, E.C., et al. (2006). Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. Neuron *52*, 255–269.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nature Genetics *39*, 61–69.

Ziller, M.J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T., et al. (2011). Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genetics *7*, e1002389.

Appendix: List of proteins identified by mass spectrometry

Appendix Table 1. All proteins identified by mass spectrometry in GFP immunoprecipitations from MeCP2-GFP and wild-type mouse brains. The proteins were GFP immunoprecipitated from MeCP2-GFP (Me2-GFP) and wild-type (Wt) mouse brains and subjected to mass spectrometry identification. All of the identified proteins are listed in an alphabetical order with their IPI database accession number and the number of peptides that identified each protein. If the protein was present only in a MeCP2-GFP purification it is marked with light blue, if present only in wild-type brain then it is marked in grey and when protein was detected in both immunoprecipitations its background is white.

Accession number	Protein description	Number of peptides Wt	Number of peptides Me2-GFP
IPI00229328	1110037F02Rik hypothetical protein LOC66185		2
IPI00131674	2210010C04Rik trypsinogen 7	1	1
IPI00121045	A2bp1 Isoform 1 of Fox-1 homolog A	5	2
IPI00121136	Acin1 Apoptotic chromatin condensation inducer Isoform 1		6
IPI00221528	Actbl2 Beta-actin-like protein 2	2	
IPI00114593	Actc1 Actin, alpha cardiac muscle 1		1
IPI00649132	Adad1 adenosine deaminase domain containing 1		1
IPI00672180	Adnp activity-dependent neurop		1
IPI00127766	Akap8l A-kinase anchor protein 8-like		7
IPI00114350	Angpt1 Angiopoietin-1		1
IPI00330263	Aqr Intron-binding protein aquarius		3
IPI00230009	Arglu1 Isoform 2 of Arginine and glutamate-rich protein 1	2	2
IPI00409462	Bat1a Spliceosome RNA helicase Bat1	2	
IPI00420725	Camk2a calcium/calmodulin-dependent protein kinase II alpha		8
IPI00875723	Camk2b 68 kDa protein		4
IPI00124695	Camk2g CamK protein kinase type II gamma chain Isoform 1		2
IPI00123755	Cbx5 Chromobox protein homolog 5	1	
IPI00153134	Ccnk Cyclin-K		1
IPI00284444	Cdc5l Cell division cycle 5-related protein		11
IPI00318428	Cenpv Isoform 1 of Centromere protein V		2
IPI00396802	Chd4 Chromodomain-helicase-DNA-binding protein 4		3
IPI00608001	Chd5 chromodomain helicase DNA binding protein 5 isoform 2		1
IPI00229604	Cherp SR-related CTD associated factor 6		16
IPI00313899	Cpne9 Copine-9	2	
IPI00314302	Cpsf2 Cleavage and polyadenylation specificity factor sub.2	1	1
IPI00421085	Cpsf6 Cleavage and polyadenylation specificity factor sub.6	1	
IPI00132376	Crnkl1 Crooked neck-like protein 1		1

IPI00120162	Csnk2a1 Casein kinase II subunit alpha		4
IPI00622716	Ctnnbl1 Beta-catenin-like protein 1		3
IPI00129146	Cugbp2 Isoform 5 of CUG-BP- and ETR-3-like factor 2		1
IPI00648836	D11Bwg0517e hexaribonucleotide binding protein 3 isoform 3	3	
IPI00894965	D11Bwg0517e Isoform 2 of Fox-1 homolog C		8
IPI00133708	D1Pas1 Putative ATP-dependent RNA helicase Pl10	1	1
IPI00396797	Ddx17 Isoform 1 of Probable ATP-dependent RNA helicase	3	11
IPI00652987	Ddx21 DEAD (Asp-Glu-Ala-Asp) box polypeptide 21		5
IPI00330406	Ddx23 DEAD (Asp-Glu-Ala-Asp) box polypeptide 23		1
IPI00468896	Ddx46 Isoform 1 of Probable ATP-dependent RNA helicase	2	
IPI00648763	Ddx5 DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	1	1
IPI00420363	Ddx5 Probable ATP-dependent RNA helicase DDX5	11	21
IPI00117771	Ddx50 ATP-dependent RNA helicase DDX50		1
IPI00130102	Des Desmin	1	
IPI00128818	Dhx15 Putative pre-mRNA-splicing factor ATP-dep. RNA helicase	1	11
IPI00265358	Dhx38 Putative uncharacterized protein		1
IPI00339468	Dhx9 Isoform 2 of ATP-dependent RNA helicase A	7	49
IPI00113635	Dkc1 H/ACA ribonucleoprotein complex subunit 4		1
IPI00307837	Eef1a1 Elongation factor 1-alpha 1	5	
IPI00119667	Eef1a2 Elongation factor 1-alpha 2		2
IPI00469260	Eftud2 116 kDa U5 small nuclear ribonucleoprotein component	1	22
IPI00126716	Eif4a3 Eukaryotic initiation factor 4A-III		12
IPI00466032	Elavl1 ELAV embryonic lethal, abnormal vision, Drosophila-like 1	4	8
IPI00410779	Elavl2 ELAV-like 2 isoform 1	2	
IPI00122451	Elavl3 Isoform HuC-L of ELAV-like protein 3		4
IPI00222990	Elavl3 Isoform HuC-S of ELAV-like protein 3	1	
IPI00466120	Elavl4 Elavl4 protein		10
IPI00119581	Fbl;LOC100044829 rRNA 2'-O-methyltransferase fibrillarin	2	1
IPI00330634	Fbll1 rRNA/tRNA 2'-O-methyltransferase fibrillarin-like protein 1		2
IPI00118205	Fhl2 Four and a half LIM domains protein 2	2	
IPI00120174	Friend of PRMT1 protein		2
IPI00117063	Fus RNA-binding protein FUS	1	1
IPI00462979	Fyttd1 Forty-two-three domain-containing protein 1, Isoform 1		4
IPI00918335	Gcfc1, GC-rich sequence DNA-binding factor 1		1
IPI00116535	Gm10120 similar to Sm D2 isoform 1		1
IPI00607069	Gm10257 hypothetical protein isoform 1		6
IPI00274175	Gm5121 similar to ribosomal protein S8 isoform 1		1
IPI00395100	Gm5409 Try10-like trypsinogen	3	
IPI00621335	Gm847 similar to hCG1643239	1	
IPI00752798	Gm9294 predicted gene 9294		1
IPI00153400	H2afj Histone H2A.J	1	5
IPI00230264	H2afx Histone H2A.x		1
IPI00137852	H2afy Isoform 1 of Core histone macro-H2A.1	2	8
IPI00652934	H2afy2 Core histone macro-H2A.2		2

IPI00331734	H2afz Histone H2A.Z		3
IPI00114232	Hdac1 Histone deacetylase 1		2
IPI00137668	Hdac2 Histone deacetylase 2		6
IPI00135456	Hdac3 histone deacetylase 3		7
IPI00228616	Hist1h1a Histone H1.1		1
IPI00223713	Hist1h1c Histone H1.2		4
IPI00111957	Hist1h2ba Histone H2B type 1-A	1	
IPI00272033	Hist2h2ac;Hist2h2ab Histone H2A type 2-C		2
IPI00229539	Hist3h2bb histone cluster 3, H2bb		1
IPI00114642	Histone H2B type 1-F/J/L		7
IPI00230730	Histone H3.2	1	1
IPI00109813	Hnrnpa0 heterogeneous nuclear ribonucleoprotein A0	5	4
IPI00553777	Hnrnpa1 Putative uncharacterized protein	3	10
IPI00828488	Hnrnpa2b1 Heterogeneous nuclear ribonucleoproteins A2/B1	4	28
IPI00269661	Hnrnpa3 Heterogeneous nuclear ribonucleoprotein A3, iso. 1	10	22
IPI00117288	Hnrnpab Heterogeneous nuclear ribonucleoprotein A/B		1
IPI00223443	Hnrnpc Heterogeneous nuclear ribonucleoproteins C1/C2, iso C1		19
IPI00130343	Hnrnpc Putative uncharacterized protein	5	1
IPI00330958	Hnrnpd Heterogeneous nuclear ribonucleoprotein D0, iso. 1		10
IPI00230086	Hnrnpd Heterogeneous nuclear ribonucleoprotein D0, iso. 2	7	
IPI00226073	Hnrnpf Isoform 1 of Heterogeneous nuclear ribonucleoprotein F	1	4
IPI00133916	Hnrnph1 Heterogeneous nuclear ribonucleoprotein H	11	15
IPI00108143	Hnrnph2, Heterogeneous nuclear ribonucleoprotein H2	7	7
IPI00816884	Hnrnph3 heterogeneous nuclear ribonucleoprotein H3		6
IPI00223253	Hnrnpk Isoform 1 of Heterogeneous nuclear ribonucleoprotein K	3	9
IPI00620362	Hnrnpl Heterogeneous nuclear ribonucleoprotein L	7	16
IPI00653643	Hnrnpl Putative uncharacterized protein		2
IPI00132443	Hnrnpm Heterogeneous nuclear ribonucleoprotein M, iso. 1	12	39
IPI00128441	Hnrnpr Putative uncharacterized protein	4	5
IPI00458583	Hnrnpu Heterogeneous nuclear ribonucleoprotein U	39	43
IPI00755892	Hnrpdl heterogeneous nuclear ribonucleoprotein D-like		7
IPI00129417	Hnrpdl Heterogeneous nuclear ribonucleoprotein D-like	3	
IPI00121760	Hnrpll Heterogeneous nuclear ribonucleoprotein L-like, iso. 1		2
IPI00323357	Hspa8 Heat shock cognate 71 kDa protei	11	21
IPI00127947	Ik IK cytokine		2
IPI00318550	Ilf2 Interleukin enhancer-binding factor 2	5	10
IPI00130591	Ilf3 Isoform 3 of Interleukin enhancer-binding factor 3	1	12
IPI00458765	Khdrbs1 KH domain-containing, RNA-binding, signal transd.ass.1	2	1
IPI00314520	Khdrbs3 KH domain-containing, RNA-binding, signal trans. ass. 3		6
IPI00462934	Khsrp Far upstream element-binding protein 2	1	
IPI00230429	Kpna3 Importin subunit alpha-3		10
IPI00129792	Kpna4 Importin subunit alpha-4		17
IPI00139301	Krt5 Keratin, type II cytoskeletal 5	4	
IPI00346834	Krt76 Keratin, type II cytoskeletal 2 oral		1

IPI00122418	Luc7l3 Isoform 2 of Luc7-like protein 3	1	
IPI00132692	Magoh-rs1 mago-nashi homolog		4
IPI00453826	Matr3 Matrin-3	9	73
IPI00115240	Mbp Isoform 1 of Myelin basic protein	1	
IPI00223377	Mbp Isoform 4 of Myelin basic protein		3
IPI00131063	Mecp2 Isoform A of Methyl-CpG-binding protein 2	1	
IPI00775806	Mecp2 Isoform B of Methyl-CpG-binding protein 2		52
IPI00228590	Mfap1a;Mfap1b Microfibrillar-associated protein 1		1
IPI00116372	Mlf2 Myeloid leukemia factor 2		6
IPI00651807	Mll2 myeloid/lymphoid or mixed-lineage leukemia 2		3
IPI00624969	Mta1 Mta1 protein		3
IPI00125745	Mta3 Isoform 1 of Metastasis-associated protein MTA3		1
IPI00408909	Mtap1a Isoform 1 of Microtubule-associated protein 1A		6
IPI00896700	Mtap1b microtubule-associated protein 1B (MAP1B)		69
IPI00154084	Myef2 Isoform 2 of Myelin expression factor 2	6	24
IPI00458056	Ncbp1 Nuclear cap-binding protein subunit 1		1
IPI00317794	Ncl Nucleolin		2
IPI00313525	Ncoa5 Nuclear receptor coactivator 5		8
IPI00551272	Ncor1 nuclear receptor co-repressor 1		20
IPI00123871	Ncor2 nuclear receptor co-repressor 2		27
IPI00554928	Nefl Neurofilament light polypeptide	5	
IPI00320016	Nono Non-POU domain-containing octamer-binding prot., iso1	3	
IPI00463468	Nop58 Nucleolar protein 58	1	
IPI00119059	Nr2f1 COUP transcription factor 1		1
IPI00263048	Numa1 nuclear mitotic apparatus protein 1		1
IPI00136169	Pabpn1 Isoform 1 of Polyadenylate-binding protein 2		1
IPI00127707	Pcbp2 Isoform 1 of Poly(rC)-binding protein 2		1
IPI00321597	Pelp1 Proline-, glutamic acid- and leucine-rich protein 1		3
IPI00133491	Phf5a PHD finger-like domain-containing protein 5A		3
IPI00331172	Plrg1 Pleiotropic regulator 1		1
IPI00317891	Pnn Pinin		3
IPI00320034	Polr2b DNA-directed RNA polymerase II subunit RPB2		1
IPI00130185	Ppp1ca Ser/thr-protein phosphatase PP1-alpha catalytic subunit		1
IPI00123862	Ppp1cc Ser/thr-protein phosphatase PP1-γ catalytic subunit	1	
IPI00117687	Pre-mRNA branch site protein p14		1
IPI00123445	Prkaa2 5'-AMP-activated protein kinase catalytic subunit alpha-2		1
IPI00480507	Prpf19 Isoform 1 of Pre-mRNA-processing factor 19		11
IPI00222760	Prpf19 Isoform 2 of Pre-mRNA-processing factor 19	2	
IPI00226155	Prpf38a Isoform 1 of Pre-mRNA-splicing factor 38A		2
IPI00458908	Prpf4 U4/U6 small nuclear ribonucleoprotein Prp4		2
IPI00320690	Prpf4b Serine/threonine-protein kinase PRP4 homolog		2
IPI00130409	Prpf6 Isoform 1 of Pre-mRNA-processing factor 6		4
IPI00121596	Prpf8 Pre-mRNA-processing-splicing factor 8	1	50
IPI00115257	Psip1 Isoform 1 of PC4 and SFRS1-interacting protein	2	

IPI00321266	Pspc1 Isoform 1 of Paraspeckle component 1	2	
IPI00751759	Ptbp2 Isoform 1 of Polypyrimidine tract-binding protein 2		4
IPI00127174	Puf60 Isoform 2 of Poly(U)-binding-splicing factor PUF60	10	2
IPI00849047	Similar to heterogeneous nuclear ribonucleoprotein U-like 2	31	39
IPI00130147	Raly Isoform 2 of RNA-binding protein Raly		8
IPI00310432	Ralyl Isoform 1 of RNA-binding Raly-like protein		5
IPI00122696	Rbbp4 Histone-binding protein RBBP4		7
IPI00122698	Rbbp7 Histone-binding protein RBBP7		1
IPI00116031	Rbm10 Isoform 1 of RNA-binding protein 10		3
IPI00471450	Rbm12b2 RNA binding motif protein 12 B2		1
IPI00404707	Rbm14 Isoform 1 of RNA-binding protein 14	6	14
IPI00170394	Rbm17 Splicing factor 45		7
IPI00221951	Rbm22 Pre-mRNA-splicing factor RBM22		2
IPI00421119	Rbm25 Isoform 1 of RNA-binding protein 25		2
IPI00127763	Rbm39 Isoform 1 of RNA-binding protein 39	12	8
IPI00177224	Rbm4b RNA-binding protein 4B		1
IPI00130160	Rbm5 Isoform 1 of RNA-binding protein 5		1
IPI00279153	Rbm6 Rbm6 protein		1
IPI00223425	Rbm9 Isoform 1 of RNA-binding protein 9		1
IPI00124979	Rbmx RNA binding motif protein, X-linked		2
IPI00663587	Rbmxrt RNA binding motif protein, X chromosome retrogene		20
IPI00122227	Rnps1 RNA/DNA-binding protein		3
IPI00221840	Rod1 Isoform 1 of Regulator of differentiation 1		1
IPI00133985	Ruvbl1 RuvB-like 1	1	2
IPI00123557	Ruvbl2 RuvB-like 2	1	1
IPI00944159	Safb scaffold attachment factor B		40
IPI00857736	Safb2 Scaffold attachment factor B2		9
IPI00458854	Sart3 Squamous cell carcinoma antigen recognized by T-cells 3		1
IPI00313539	Scai Isoform 1 of Protein SCAI	2	1
IPI00408796	Sf3a1 Splicing factor 3 subunit 1	2	14
IPI00380281	Sf3a2 splicing factor 3a, subunit 2		4
IPI00137848	Sf3a3 Splicing factor 3A subunit 3	1	10
IPI00623284	Sf3b1 Splicing factor 3B subunit 1	2	41
IPI00349401	Sf3b2 splicing factor 3b, subunit 2	1	16
IPI00122011	Sf3b3 Isoform 1 of Splicing factor 3B subunit 3	4	26
IPI00154082	Sf3b4 Splicing factor 3B subunit 4		1
IPI00420807	Sfrs1;LOC100048559 Splicing factor, arginine/serine-rich 1, iso. 1	2	11
IPI00468994	Sfrs11 splicing factor, arginine/serine-rich 11 isoform 1	2	1
IPI00117232	Sfrs13a Isoform 3 of Splicing factor, arginine/serine-rich 13A		2
IPI00405760	Sfrs14 Putative splicing factor, arginine/serine-rich 14	17	66
IPI00121135	Sfrs2 Splicing factor, arginine/serine-rich 2		1
IPI00129323	Sfrs3 Isoform Long of Splicing factor, arginine/serine-rich 3	3	3
IPI00607076	Sfrs4 splicing factor, arginine/serine-rich 4		1
IPI00314709	Sfrs5 Splicing factor, arginine/serine-rich 5		2

IPI00310880	Sfrs6 arginine/serine-rich splicing factor 6		5
IPI00153743	Sfrs7 Isoform 2 of Splicing factor, arginine/serine-rich 7		7
IPI00117932	Sin3a Isoform 1 of Paired amphipathic helix protein Sin3a		1
IPI00229571	Sltm Isoform 1 of SAFB-like transcription modulator		21
IPI00314654	Smarca1 Probable global transcription activator SNF2L1, iso. 2		1
IPI00381019	Smarcc2 Isoform 2 of SWI/SNF complex subunit SMARCC2		6
IPI00331342	Smu1 Isoform 1 of WD40 repeat-containing protein SMU1		4
IPI00420329	Snrnp200 Small nuclear ribonucleoprotein 200	2	43
IPI00461621	Snrnp40 U5 small nuclear ribonucleoprotein 40 kDa protein		7
IPI00625105	Snrnp70 Isoform 1 of U1 small nuclear ribonucleoprotein 70 kDa		7
IPI00170008	Snrpa1 U2 small nuclear ribonucleoprotein A'	11	12
IPI00114052	Snrpb Small nuclear ribonucleoprotein-associated protein B		3
IPI00322749	Snrpd1 Small nuclear ribonucleoprotein Sm D1	1	3
IPI00119224	Snrpd3 Small nuclear ribonucleoprotein Sm D3		3
IPI00133955	Snrpe Small nuclear ribonucleoprotein E		1
IPI00317298	Snw1 SNW domain-containing protein 1		5
IPI00134457	Son Isoform 1 of Protein SON		1
IPI00118438	Srrm1 Isoform 2 of Serine/arginine repetitive matrix protein 1		1
IPI00225062	Srrm2 Isoform 3 of Serine/arginine repetitive matrix protein 2		12
IPI00134300	Ssb Lupus La protein homolog		1
IPI00120344	Supt16h FACT complex subunit SPT16		1
IPI00406117	Syncrip Heterogeneous nuclear ribonucleoprotein Q, isoform 1		2
IPI00127982	Sytl4 Isoform 1 of Synaptotagmin-like protein 4		1
IPI00121758	Tardbp TAR DNA-binding protein 43		3
IPI00223056	Tbl1x F-box-like/WD repeat-containing protein TBL1X		7
IPI00308283	Tbl1xr1 F-box-like/WD repeat-containing protein TBL1XR1		16
IPI00458057	Tex10 Tex10 protein		1
IPI00112101	Tfip11 Tuftelin-interacting protein 11		1
IPI00556768	Thrap3 Thyroid hormone receptor-associated protein 3		1
IPI00377298	Tra2a transformer-2 alpha		3
IPI00139259	Tra2b Isoform 1 of Transformer-2 protein homolog beta		7
IPI00312128	Trim28 Isoform 1 of Transcription intermediary factor 1-beta	7	13
IPI00229801	Trp53bp1 transformation related protein 53 binding protein 1		32
IPI00110753	Tuba1a Tubulin alpha-1A chain		2
IPI00117348	Tuba1b Tubulin alpha-1B chain	4	
IPI00109061	Tubb2b Tubulin beta-2B chain	4	1
IPI00411097	U2 snRNP-associated SURP motif-containing protein isoform 2		18
IPI00318548	U2af1 Splicing factor U2AF 35 kDa subunit	3	5
IPI00221628	U2af1l4 Splicing factor U2AF 26 kDa subunit	1	
IPI00113746	U2af2 Splicing factor U2AF 65 kDa subunit	8	7
IPI00138892	Uba52, Ubiquitin-60S ribosomal protein L40	1	3
IPI00329998	Uncharacterised 11 kDa protein	6	9
IPI00474144	Uncharacterised 42 kDa protein	3	
IPI00123333	Wbp11 WW domain-binding protein 11		5

IPI00136252	Wdr18 WD repeat-containing protein 18		4
IPI00139957	Wdr5 WD repeat-containing protein 5		1
IPI00120744	Xab2 Pre-mRNA-splicing factor SYF1		2
IPI00653196	Yap1, Putative uncharacterized protein		1
IPI00621973	Ylpm1 YLP motif containing 1		9
IPI00464155	Zbtb4 zinc finger and BTB domain containing 4	1	
IPI00121264	Zfml Isoform 4 of Zinc finger protein 638		2
IPI00314507	Zfp326 Isoform 1 of Zinc finger protein 326	2	17
IPI00330097	Zfp871 zinc finger protein 871		1
IPI00918528	Zfr Protein		1
IPI00131810	Zfr Zinc finger RNA-binding protein	1	
IPI00762823	Zfr2 zinc finger RNA binding protein 2	1	7

Appendix Table 2. All proteins identified as specific interaction partners of MeCP2.
The list of proteins identified by mass spectrometry in GFP immunoprecipitations from MeCP2-GFP and wild-type mouse brains (Exp. #1; Appendix Table 1) was compared to a similar list from a biological replicate of the experiment (Exp. #2; M. Lyst). A protein was considered as a specific interaction partner of MeCP2 only if it was represented by at least 2 peptides in both purifications from MeCP2-GFP brains and no peptides were present in both control purifications from wild-type brains.

Number of peptides			
Exp. #1	Exp. #2	Protein symbol	Protein name
52	57	Mecp2	Isoform B of Methyl-CpG-binding protein 2
27	5	Ncor2	Nuclear receptor co-repressor 2
20	9	Ncor1	Nuclear receptor co-repressor 1
17	13	Kpna4	Importin subunit alpha-4
16	12	Tbl1xr1	F-box-like/WD repeat-containing protein TBL1XR1
12	15	Eif4a3	Eukaryotic initiation factor 4A-III
11	19	Prpf19	Pre-mRNA-processing factor 19, Isoform 1
11	7	Cdc5l	Cell division cycle 5-related protein
10	7	Kpna3	Importin subunit alpha-3
8	3	Camk2a	Calcium/calmodulin-dependent protein kinase II alpha
8	2	Ncoa5	Nuclear receptor coactivator 5
8	2	Raly	RNA-binding protein Raly, Isoform 2
7	10	Rbbp4	Histone-binding protein RBBP4
7	10	Tra2b	Transformer-2 protein homolog beta, Isoform 1
7	7	Sfrs7	Splicing factor, arginine/serine-rich 7, Isoform 2
7	6	Snrnp70	U1 small nuclear ribonucleoprotein 70 kDa, Isoform 1
7	5	Hdac3	Histone deacetylase 3

7	5	Tbl1x	F-box-like/WD repeat-containing protein TBL1X
6	17	Smarcc2	SWI/SNF complex subunit SMARCC2, Isoform 2
5	2	Snw1	SNW domain-containing protein 1
5	2	Wbp11	WW domain-binding protein 11
4	9	Csnk2a1	Casein kinase II subunit alpha
4	9	Wdr18	WD repeat-containing protein 18
4	5	Magoh-rs1	Mago-nashi homolog
4	5	Smu1	WD40 repeat-containing protein SMU1, Isoform 1
4	4	Fyttd1	Forty-two-three domain-containing protein 1, iso. 1
3	58	Chd4	Chromodomain-helicase-DNA-binding protein 4
3	27	Mta1	Metastasis-associated protein Mta1
3	12	Pelp1	Proline-, glutamic acid- and leucine-rich protein 1
3	7	Pnn	Pinin
3	5	Tra2a	Transformer-2 alpha
3	3	Aqr	Intron-binding protein aquarius
3	3	H2afz	Histone H2A.Z
3	3	Rnps1	RNA/DNA-binding protein
2	8	Hist2h2ac	Histone H2A type 2-C
2	6	H2afy2	Core histone macro-H2A.2
2	5	Hdac1	Histone deacetylase 1
2	5	Sfrs5	Splicing factor, arginine/serine-rich 5
2	4	Tuba1a	Tubulin alpha-1A chain
2	2	Ik	IK cytokine